

# Polyhedra and Optimization in Connection with a Weak Majorization Ordering

Geir Dahl

Preprint 10

December 12, 1994





# Polyhedra and Optimization in Connection with a Weak Majorization Ordering

Geir Dahl \*

December 12, 1994

## Abstract

We introduce the concept of weak  $k$ -majorization extending the classical notion of weak sub-majorization. For integers  $k$  and  $n$  with  $k \leq n$  a vector  $x \in \mathbb{R}^n$  is weakly  $k$ -majorized by a vector  $q \in \mathbb{R}^k$  if the sum of the  $r$  largest components of  $x$  does not exceed the sum of the  $r$  largest components of  $q$ , for  $r = 1, \dots, k$ . For a given  $q$  the set of vectors weakly  $k$ -majorized by  $q$  defines a polyhedron  $P(q; k)$ , and we determine all its vertices. We also determine the vertices and a complete and nonredundant linear description of the integer hull of  $P(q; k)$ . The results are used to give simple and efficient (polynomial time) algorithms for associated linear and integer linear programming problems.

*Keywords:* Majorization; polyhedral combinatorics.

## 1 Introduction

The concept of *majorization* has been investigated in a number of branches of mathematics and statistics. The notion reflects to what extent components of vectors are “spread out”. For  $p, q \in \mathbb{R}^n$  one says that  $p$  is *weakly sub-majorized* by  $q$  if  $\sum_{j=1}^r p_{[j]} \leq \sum_{j=1}^r q_{[j]}$  for  $r = 1, \dots, n$ . Here  $p_{[j]}$  denotes the  $j$ 'th largest component of  $p$ . If also  $\sum_{j=1}^n p_j = \sum_{j=1}^n q_j$  holds,  $p$  is *majorized* by  $q$  and we write  $p \prec q$ . Several equivalent conditions for (weak sub-) majorization are known (see [8]); for instance,  $p \prec q$  iff there is a doubly stochastic matrix  $M \in \mathbb{R}^{n,n}$  (i.e.  $M$  has nonnegative elements and all row and column sums are 1) with  $p = Mq$ . This result can be combined with a theorem of Birkhoff and von Neumann to give the following geometric interpretation of majorization:  $p \prec q$  if and only if  $p$  lies in the convex hull of the set of vectors obtained by permuting the components of  $q$ . The Birkhoff-von Neumann theorem says that the set of doubly stochastic matrices in  $\mathbb{R}^{n,n}$  is the convex hull of the set of

---

\*Institute of Informatics, University of Oslo, P.O.Box 1080, Blindern, 0316 Oslo, Norway  
(Email:geird@ifi.uio.no)

all permutation matrices, i.e.  $n \times n$  matrices where each row and each column consists of all zeros except for one entry which is 1. A similar characterization holds for weak submajorization, see [8].

An extensive treatment of the theory of majorization as well as its applications is given in the book by Marshall and Olkin [8]. An example of majorization in combinatorial analysis is the theorem of Gale and Ryser. It characterizes those vectors  $r$  and  $c$  for which there is a 0/1 matrix with row sums  $r$  (i.e.,  $i$ 'th row sum is  $r_i$ ) and column sums  $c$ ; the condition is a simple majorization condition in terms of  $c$  and  $r$ . A relation between edge coloring and majorization is discussed in [4]. Majorization also arises in e.g. matrix theory (Hadamard type inequalities, singular values etc.), numerical analysis (condition numbers, matrix approximation etc.) as well as in probability theory and statistics (reliability, stochastic ordering etc.), see [8].

For generalizations of majorization within a measure theoretical framework as well as statistical interpretations, see the extensive treatment in [11]. In [2] approximate majorization is studied.

A central concept in the theory of majorization is that of a Schur-convex function. A function  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  that preserves the ordering given by majorization is called *Schur-convex*, thus  $\phi(x) \leq \phi(q)$  whenever  $x \prec q$ . Therefore  $q$  maximizes  $\phi(x)$  over the set  $x \prec q$ , so maximization of Schur-convex functions is easy. The following general method has produced interesting inequalities in the areas mentioned above, see [8]: find some (simple) underlying majorization which combined with a suitable Schur-convex function leads to the inequality. A simple example, which is useful in this work, is the *rearrangement inequality* due to Hardy, Littlewood and Polya, see [7], [8]. Let  $a_1, \dots, a_n$  and  $b_1, \dots, b_n$  be real numbers. Then we have:

$$\sum_{i=1}^n a_{[i]} b_{[n-i+1]} \leq \sum_{i=1}^n a_i b_i \leq \sum_{i=1}^n a_{[i]} b_{[i]}. \quad (1)$$

In this paper we study a majorization concept derived from weak submajorization by requiring that the ordering of the  $r$  largest components of vectors only hold for  $r \leq k$  where  $k$  is some number not exceeding  $n$ . This concept, called *weak  $k$ -majorization*, is therefore a weaker notion than weak sub-majorization and it measures the “concentration” of sums of subvectors up to a given size. Unlike for Schur-convex functions, optimization with a weak  $k$ -majorization constraint is not a trivial task, and a main purpose of this paper is to study such optimization problems.

This paper is organized as follows. In section 2 we introduce weak  $k$ -majorization and describe some of its basic properties. We introduce linear programming (LP) problems with a weak  $k$ -majorization constraint and consider the special case  $k = n$  in some detail. In section 3 we study, for  $k < n$ ,  $k$ -majorization polyhedra being the feasible set of these LP problems, and determine all the vertices of these. This also leads to a simple (polynomial) algorithm for solving such LP problems. The remaining part of the paper treats integer linear programming (ILP) problems over the mentioned polyhedra, and vertices as well as complete linear descriptions

are given for the integer hull of  $k$ -majorization polyhedra. Also a linear time algorithm for solving integer linear programming problems with a weak  $k$ -majorization constraint is given.

**Notation.**  $\mathbb{R}$ ,  $\mathbb{Z}$  and  $\mathbb{Q}$  denote the set of real, integral and rational numbers, respectively, and  $\mathbb{R}^{m,n}$  is the set of real  $m \times n$  matrices. For  $q \in \mathbb{R}^k$  we define the *tail average*  $\bar{q}_{s,k} := (1/(k-s+1)) \sum_{j=s}^k q_j$ . For each positive integer  $t$  we let  $\mathbb{N}_t := \{1, \dots, t\}$ , and for  $x \in \mathbb{R}^n$ ,  $x_{[j]}$  is the  $j$ 'th largest component of  $x$ . When  $S \subseteq \mathbb{N}_n$  we let  $|S|$  denote the cardinality of  $S$ , and if  $x \in \mathbb{R}^n$  we define  $x(S) := \sum_{j \in S} x_j$ . For concepts and results concerning polyhedra and linear inequalities, see [10]. The characteristic cone of a polyhedron  $P$  is denoted by  $\text{char.cone}(P)$ . A pointed polyhedron is *integral* if all its vertices are integral. The *integer hull*  $P_I$  of a (rational) polyhedron  $P$  is the polyhedron being the convex hull of all the integral points in  $P$ . A set  $A \subseteq \mathbb{R}^n$  is *symmetric* if  $\pi x \in A$  whenever  $x \in A$  and  $\pi$  is a permutation on  $\mathbb{N}_n$  (i.e. a bijection) and where we define  $\pi x = (x_{\pi(1)}, \dots, x_{\pi(n)})$ . For integers  $r, s$  with  $r \leq s$  we let  $C_{r,s} \in \mathbb{R}^{s,s}$  denote the *0/1-circulant matrix* with elements 0 and 1 and where the ones in row  $i$  are in the positions  $j \equiv i, i+1, \dots, i+(r-1) \pmod{s}$  (so  $r$  consecutive ones are shifted one position to the right for each row). We let  $e_i \in \mathbb{R}^n$  be the  $i$ 'th unit (coordinate) vector in  $\mathbb{R}^n$ , i.e., the  $i$ 'th component of  $e_i$  is 1 and all other components are 0.

## 2 Weak $k$ -majorization and optimization

We study basic properties of weak  $k$ -majorization. Associated optimization problems and polyhedra are also introduced, and the special case  $k = n$  is treated in some detail.

Let, throughout this section,  $k$  and  $n$  be two given integers such that  $k \leq n$ , and let  $q \in \mathbb{R}^k$  be a given vector called the *majorant*. We extend the concept of weak sub-majorization into  $k$ -majorization as follows. We say that  $x \in \mathbb{R}^n$  is *weakly  $k$ -majorized* by  $q$  and write  $p \prec_k q$  if the following conditions hold:

$$\sum_{j=1}^r x_{[j]} \leq \sum_{j=1}^r q_{[j]} \quad \text{for all } r \in \mathbb{N}_k. \quad (2)$$

The order of the components in the vectors  $x$  and  $q$  is irrelevant in this definition, so if  $x \prec_k q$ , then also  $x' \prec_k q'$  where  $x'$  and  $q'$  are permutations of  $x$  and  $q$ , respectively. We see that  $x \prec_1 q$  iff  $\max_j x_j \leq \max_j q_j$ , and that  $x \prec_n q$  means that  $x$  is weakly sub-majorized by  $q$ . We also see that, for general  $k$ ,  $x \prec_k q$  iff  $x^{[k]}$  is weakly sub-majorized by  $q^{[k]}$  where  $x^{[k]} = (x_{[1]}, \dots, x_{[k]})$  and  $q^{[k]} = (q_{[1]}, \dots, q_{[k]})$ . Thus weak  $k$ -majorization corresponds to weak sub-majorization on the vectors consisting of the  $k$  largest components. Therefore one can formulate equivalent conditions for  $x \prec_k q$  in terms of  $x^{[k]}$  and  $q^{[k]}$ . For instance,  $x \prec_k q$  if and only if there is a doubly stochastic matrix  $M \in \mathbb{R}^{k,k}$  such that  $x^{[k]} = q^{[k]} M$ .

For  $z \in \mathbb{R}^n$  we define the real-valued function  $L_z$  on the interval  $[0, 1]$  by (i)  $L_z(r/k) = \sum_{j=1}^r z_{[j]}$  for  $r = 0, \dots, k$ , and (ii)  $L_z$  is linear

on each subinterval  $[r/k, (r+1)/k]$ , for  $r = 0, \dots, k-1$ . This function is piecewise linear, continuous and concave, and it satisfies  $L_z(0) = 0$ ,  $L_z(1) = \sum_{j=1}^k z_{[j]}$ . We call  $L_z$  the *L-function* associated with  $z$ . The dependency on  $k$  is suppressed in the notation  $L_z$  as  $k$  is clear from the context. We mention that in [8] Lorentz-curves are discussed, they are similar to our *L-curves* defined below but with “opposite” ordering of the components. The *L-curve* associated with  $z$  is the graph of  $L_z$ . Majorization concepts may be described in terms of *L-functions* as follows:  $x \prec_k q$  if and only if  $L_x \leq L_q$  (with componentwise ordering), i.e. the *L-curve* associated with  $x$  lies below the *L-curve* associated with  $q$ .

Our main interest is to study optimization problems with a majorization constraint and with a linear (and nonsymmetric) objective function. The objective function in our different problems will be  $c^T x = \sum_{j=1}^n c_j x_j$  which is to be maximized over all vectors  $x \in \mathbb{R}^n$  (or  $x \in \mathbb{Z}^n$ ) that satisfy a  $k$ -majorization constraint. Thus, for instance, we shall study the set of points in  $\mathbb{R}^n$  that are weakly  $k$ -majorized by some vector  $q$ .

We assume hereafter that  $q$  satisfies  $q_1 \geq \dots \geq q_k \geq 0$ . Let  $c \in \mathbb{R}^n$  be a nonnegative objective function. Consider the following optimization problem:

$$\max \{c^T x \mid x \prec_k q\}. \quad (3)$$

A possible interpretation of this problem is as follows. Consider  $n$  projects, numbered from 1 to  $n$ , and assume that  $c_j$  is an “expected value” or profit associated with project  $j$ . Let the variable  $x_j$  represent the investment in project  $j$ . To keep down the overall risk of investing in failure projects, one may require that the investments are “suitably spread out”. This requirement may be implemented by saying that the investment vector  $x = (x_1, \dots, x_n)$  is weakly  $k$ -majorized by a given vector  $q$ . This prevents investing too much in a single project, too much in any pair of projects etc. The risk aversion may be reflected in the choice of  $k$  and  $q$ .

Let  $P(q; k)$  denote the set of feasible solutions in the optimization problem (3), so  $P(q; k) = \{x \in \mathbb{R}^n \mid x \prec_k q\}$ . Note that  $x \prec_k q$  if and only if  $x(S) \leq q(\mathbb{N}_r)$  for each subset  $S$  of  $\mathbb{N}_n$  with  $|S| = r \leq k$  because the maximum value of  $x(S)$  taken over all such subsets is  $\sum_{j=1}^r x_{[j]}$  (confer the rearrangement inequality). It follows that the set  $P(q; k)$  is a polyhedron, in fact

$$P(q; k) = \{x \in \mathbb{R}^n \mid x(S) \leq q(\mathbb{N}_r) \text{ for all } S \subseteq \mathbb{N}_n \text{ with } r = |S| \leq k\}. \quad (4)$$

The polyhedron  $P(q; k)$ , called a *k-majorization polyhedron*, is unbounded, its characteristic cone is  $-\mathbb{R}^n$  and it is pointed, i.e. its minimal faces are vertices. Furthermore,  $P(q; k)$  is symmetric.

Each  $n$ -majorization polyhedron may be viewed as a *polymatroid* (see e.g. [5], [6]). Define the set function  $f(S) = \sum_{j=1}^r q_j$  for each  $S \subseteq \mathbb{N}_n$  where  $r := |S|$ . Then it follows from the assumption  $q_1 \geq \dots \geq q_k \geq 0$  that: (i)  $f$  is monotone, i.e.,  $f(S) \leq f(T)$  for  $S \subseteq T \subseteq \mathbb{N}_n$ , (ii)  $f$  is submodular, i.e.,  $f(S \cup T) + f(S \cap T) \leq f(S) + f(T)$  for all  $S, T \subseteq \mathbb{N}_n$ , and (iii)  $f$  is nonnegative. From (4) we see that  $P(q; n) = \{x \in \mathbb{R}^n \mid$

$x(S) \leq f(S)$  for all  $S \subseteq \mathbb{N}_n$ , so this is the polymatroid associated with the submodular function  $f$ . Edmonds (see [3]) showed that maximizing a nonnegative linear objective function over a polymatroid can be done using the greedy algorithm. Thus, in the case  $k = n$ , we can solve the problem (3) by the following greedy algorithm (assume for simplicity that  $c_1 \geq \dots \geq c_n \geq 0$ ):  $x_1 = f(\{1\}) = q_1$ ,  $x_2 = f(\{1, 2\}) - f(\{1\}) = (q_1 + q_2) - q_1 = q_2$ ,  $\dots$ ,  $x_n = f(\mathbb{N}_n) - f(\mathbb{N}_{n-1}) = q_n$ . The optimality of this solution could also be seen as a direct consequence of the rearrangement inequality (1). For  $k < n$ , however,  $P(q; k)$  may not be a polymatroid and therefore the greedy solution which is  $x_j = q_j$  for  $j \leq k$  and  $x_j = q_k$  for  $j > k$  may not be optimal in (3). As a simple example, let  $n = 3$ ,  $k = 2$ ,  $q = (2, 1)$  and  $c = (1, 1, 1)$ . The greedy algorithm produces the nonoptimal solution  $(2, 1, 1)$  while the optimal solution is  $(3/2, 3/2, 3/2)$ . In the next section we shall explain how to “repair” this by giving an algorithm that solves (3) for general  $k$ . This development will be done by studying properties of the  $k$ -majorization polyhedra.

It turns out that  $n$ -majorization polyhedra have a simple, but interesting combinatorial structure as briefly discussed next. Consider the face  $M(q; n)$  of  $P(q; n)$  induced by the valid inequality  $x(\mathbb{N}_n) \leq q(\mathbb{N}_n)$ , i.e.  $M(q; n) = \{x \in P(q; n) \mid x(\mathbb{N}_n) = q(\mathbb{N}_n)\}$ . Then  $M(q; n)$  consists of all vectors  $x$  that are majorized by  $q$  and it is the base polyhedron of the polymatroid  $P(q; n)$ .  $M(q; k)$  is bounded polyhedron, i.e. a polytope, and its extreme points are all the  $n!$  permutations of  $q$ . Assuming that  $q_1 > \dots > q_n > 0$ , it can be shown that  $\dim(M(q; k)) = n - 1$  and that the linear system  $x(\mathbb{N}_n) = \sum_{j=1}^n q_j$ ,  $x(S) \leq \sum_{j=1}^r q_j$  for all  $S \subseteq \mathbb{N}_n$ ,  $r := |S|$  gives a complete and nonredundant linear description of  $M(q; n)$ . Two vertices are adjacent on  $M(q; n)$  (i.e., there is a 1-dimensional face containing both vertices) if and only if the permutations (of  $q$ ) defining the two vertices differ by an adjacent transposition. Finally,  $M(q; n)$  is *simple*, meaning that each vertex is contained in (the minimum number of)  $n$  facets. An interesting special case is when  $q = (n, n-1, \dots, 1)$ . Then  $M(q; n)$  is called the *permutahedron*, see [1], [12]. All the properties described above for  $M(q; n)$  are well known to hold for the permutahedron. In fact, one can show that (for  $q_1 > \dots > q_n > 0$ )  $M(q; n)$  and the permutahedron are combinatorially equivalent.

The remaining part of this paper is devoted to the study of  $P(q; k)$  whenever  $k < n$ .

### 3 Vertices of $k$ -majorization polyhedra

In this section we find all the vertices of the  $k$ -majorization polyhedron  $P(q; k)$ . Throughout this section we assume that  $k < n$  and also (to avoid degenerate situations) that the given majorant  $q \in \mathbb{R}^k$  satisfies  $q_1 > \dots > q_k > 0$ . We define  $Q = \sum_{j=1}^k q_j$ .

Define, for  $s = 0, \dots, k-1$ , the vector  $w^s$ , called the  $s$ 'th  $q$ -average by

$$w^s := (q_1, \dots, q_s, \bar{q}_{s+1:k}, \dots, \bar{q}_{s+1:k}) \quad (5)$$

Note the special cases  $w^0 = (\bar{q}_{1:k}, \dots, \bar{q}_{1:k})$  and  $w^{k-1} = (q_1, q_2, \dots, q_k, \dots, q_k)$  and that  $\sum_{j=1}^k w_j^s = Q$  as we have  $\sum_{j=1}^s q_j + (k-s)\bar{q}_{s+1:k} = \sum_{j=1}^k q_j = Q$ . We also observe that  $w^s \prec_k q$ . In fact, the  $L$ -functions  $L_{w^s}$  and  $L_q$  (associated with  $w^s$  and  $q$ , respectively) coincide on  $[0, s/k]$  and  $L_{w^s}(1) = L_q(1)$ . Since  $L_{w^s}$  is linear on  $[s/k, 1]$  and  $L_q$  is concave, it follows that  $L_{w^s} \leq L_q$  and therefore  $w^s \prec_k q$ . Note that, among the vectors that are weakly  $k$ -majorized by  $q$ , the  $q$ -average  $w^s$  is maximally spread out in its first  $k$  components and minimally spread out (constant) in its remaining components.

The main result of this section says that the vertices of  $P(q; k)$  are the permutations of the  $q$ -averages. In order to prove this result we shall need a lemma on tail averages which will be of use later as well.

**Lemma 1** *Let  $c = (c_1, \dots, c_k) \in \mathbb{R}^k$  satisfy  $c_1 \geq \dots \geq c_{k-1}$  (note that  $c_k$  may be arbitrary). Then there is an  $m \in \{1, \dots, k\}$  such that*

$$\bar{c}_{1:k} \geq \dots \geq \bar{c}_{m:k} \leq \bar{c}_{m+1:k} \leq \dots \leq \bar{c}_{k:k} = c_k. \quad (6)$$

**Proof.** For  $t = 1, \dots, k-1$  the following relation holds between consecutive tail averages:  $\bar{c}_{t:k} = c_t/(k-t+1) + \bar{c}_{t+1:k}(k-t)/(k-t+1)$ . Thus  $\bar{c}_{t:k}$  is a convex combination of  $c_t$  and  $\bar{c}_{t+1:k}$  which leads to, for  $t = 1, \dots, k$ ,

$$\begin{aligned} \text{(i)} \quad c_t \leq \bar{c}_{t+1:k} &\Rightarrow c_t \leq \bar{c}_{t:k} \leq \bar{c}_{t+1:k}; \\ \text{(ii)} \quad c_t \geq \bar{c}_{t+1:k} &\Rightarrow c_t \geq \bar{c}_{t:k} \geq \bar{c}_{t+1:k}. \end{aligned} \quad (7)$$

Let  $t = k-1$ . While  $c_t \leq \bar{c}_{t+1:k}$  and  $t \geq 1$  set  $t := t-1$  and repeat this procedure. Assume that we here terminate with  $t = t'$ . If  $t' = 1$ , (7)(i) gives  $\bar{c}_{1:k} \leq \dots \leq \bar{c}_{k:k}$ , so (6) holds with  $m = 1$ . If  $t' > 1$ , we have (again by (7)(i))  $c_{t'} > \bar{c}_{t'+1:k} \leq \dots \leq \bar{c}_{k:k}$ , and then (7)(ii) implies that  $\bar{c}_{t':k} \geq \bar{c}_{t'+1:k}$ . Since  $c$  is nonincreasing, we can repeatedly use (7)(ii) and thereby obtain (6) with  $m = t' + 1$ .  $\square$

**Theorem 2** *The vertex set of  $P(q; k)$  is the set  $W(q; k)$  of vectors that can be obtained as permutations of some  $w^s$  for  $s \in \{0, \dots, k-1\}$ .*

**Proof.** We first show that each vertex of  $P(q; k)$  lies in  $W(q; k)$ . by proving that each bounded LP problem over  $P(q; k)$  has an optimal solution lying in  $W(q; k)$ . Let  $d \in \mathbb{R}^n$  be an objective function and consider the LP problem  $\max\{d^T x \mid x \in P(q; k)\}$ . If some  $d_j < 0$ , then the problem is unbounded as  $\lambda e_j \in P(q; k)$  for each  $\lambda \geq 0$ . We may therefore assume that  $d_1 \geq \dots \geq d_n \geq 0$ . From the rearrangement inequality (1) and the fact that only the  $k$  largest components of  $x$  determine the feasibility of a point, we may assume that  $x_1 \geq \dots \geq x_k = \dots = x_n$ . Thus our LP



problem reduces another LP problem

$$\begin{aligned}
& \max \quad \sum_{j=1}^k c_j x_j \\
& \text{subject to} \\
& \text{(i)} \quad x(\mathbb{N}_r) \leq q(\mathbb{N}_r) \quad \text{for } r = 1, \dots, k; \\
& \text{(ii)} \quad x_{j+1} - x_j \leq 0 \quad \text{for } j = 1, \dots, k-1;
\end{aligned} \tag{8}$$

where we define  $c_j = d_j$  for  $j = 1, \dots, k-1$  and  $c_k = \sum_{j=k}^n d_j$ . Note that we here have  $c_1 \geq \dots \geq c_{k-1}$  while  $c_k$  can have “any” value (not exceeding  $(n-k+1)d_k$ ). The LP dual of this problem is

$$\begin{aligned}
& \min \quad \sum_{i=1}^k y_i \sum_{j=1}^i q_j \\
& \text{subject to} \\
& \text{(i)} \quad \sum_{i=r}^k y_i + z_{r-1} - z_r = c_r \quad \text{for } r = 1, \dots, k; \\
& \text{(ii)} \quad y_1, \dots, y_k, z_1, \dots, z_{k-1} \geq 0 \quad (\text{and } z_0 = z_k = 0).
\end{aligned} \tag{9}$$

The variables  $z_0$  and  $z_k$  are used for notational simplicity only; these are not variables in the dual problem. We now apply Lemma 1 to the objective function  $c$  and let  $m$  be such that the inequalities (6) hold. Let  $s = m-1$ , and observe that  $s \in \{0, \dots, k-1\}$ . Let  $\hat{x} \in \mathbb{R}^k$  and  $(\hat{y}, \hat{z}) \in \mathbb{R}^k \times \mathbb{R}^{k-1}$  be given by

$$\begin{aligned}
\text{(i)} \quad \hat{x}_j &= \begin{cases} q_j & \text{for } j = 1, \dots, s, \\ \bar{q}_{s+1:k} & \text{for } j = s+1, \dots, k; \end{cases} \\
\text{(ii)} \quad \hat{y}_j &= \begin{cases} c_j - c_{j+1} & \text{for } j = 1, \dots, s-1, \\ c_s - \bar{c}_{s+1:k} & \text{for } j = s, \\ 0 & \text{for } j = s+1, \dots, k-1, \\ \bar{c}_{s+1:k} & \text{for } j = k; \end{cases} \\
\text{(iii)} \quad \hat{z}_j &= \begin{cases} 0 & \text{for } j = 1, \dots, s, \\ (k-j)(\bar{c}_{j+1:k} - \bar{c}_{s+1:k}) & \text{for } j = s+1, \dots, k-1. \end{cases}
\end{aligned}$$

The solution  $\hat{x}$  can be viewed as the projection of  $w^s$  into the space of the first  $k$  variables, which implies that  $\hat{x}$  is feasible in (8). We next show that  $(\hat{y}, \hat{z})$  is feasible in the dual problem (9). Clearly,  $\hat{y}_j \geq 0$  for  $j < s$  since  $c_j \geq c_{j+1}$ . Also,  $\hat{y}_s = c_s - \bar{c}_{s+1:k} \geq 0$  according to our choice of  $s$ ; see Lemma 1 and (7). The remaining components of  $\hat{y}$  are trivially nonnegative, and therefore  $\hat{y} \geq 0$ . It follows from Lemma 1 that  $\hat{z} \geq 0$ .

It remains to verify that  $(\hat{y}, \hat{z})$  satisfies constraint (9)(i). If  $r \geq s+2$ , we have  $\sum_{i=r}^k \hat{y}_i + \hat{z}_{r-1} - \hat{z}_r = \bar{c}_{s+1:k} + (k-r+1)(\bar{c}_{r:k} - \bar{c}_{s+1:k}) - (k-r)(\bar{c}_{r+1:k} - \bar{c}_{s+1:k}) = (k-r+1)\bar{c}_{r:k} - (k-r)\bar{c}_{r+1:k} = c_r$ . Next, if  $r = s+1$ , we obtain  $\sum_{i=s+1}^k \hat{y}_i + \hat{z}_s - \hat{z}_{s+1} = \bar{c}_{s+1:k} - (k-s-1)(\bar{c}_{s+2:k} - \bar{c}_{s+1:k}) = (k-s)\bar{c}_{s+1:k} - (k-s-1)\bar{c}_{s+2:k} = c_{s+1}$ . Finally, when  $r \leq s$  we have  $\sum_{i=r}^k \hat{y}_i + \hat{z}_{r-1} - \hat{z}_r = \sum_{i=r}^{s-1} (c_i - c_{i+1}) + c_s - \bar{c}_{s+1:k} + \bar{c}_{s+1:k} = (c_r - c_{r+1}) + (c_{r+1} - c_{r+2}) + \dots + (c_{s-1} - c_s) + c_s = c_r$ . This proves that  $(\hat{y}, \hat{z})$  is dual feasible.

The value of the (primal) objective function of (8) in the point  $\hat{x}$  is  $\sum_{j=1}^k c_j \hat{x}_j = \sum_{i=1}^s c_i q_i + (1/(k-s)) \sum_{i=s+1}^k q_i \sum_{j=s+1}^k c_j = \sum_{i=1}^s c_i q_i + \bar{c}_{s+1:k} \sum_{j=s+1}^k q_j$ . The value of the (dual) objective function of (9) in the point  $(\hat{y}, \hat{z})$  is  $\sum_{i=1}^k \hat{y}_i \sum_{j=1}^i q_j = \sum_{i=1}^{s-1} (c_i - c_{i+1}) \sum_{j=1}^i q_j + (c_s - \bar{c}_{s+1:k}) \sum_{j=1}^s q_j + \bar{c}_{s+1:k} \sum_{j=1}^k q_j = \sum_{j=1}^s c_j q_j + \bar{c}_{s+1:k} \sum_{j=s+1}^k q_j$ . We see that these two values coincide. Therefore, by linear programming duality, it follows that  $\hat{x}$  is optimal in the primal problem (8) and  $(\hat{y}, \hat{z})$  is optimal in (9). It follows from the initial problem reduction that  $w^s$  is an optimal solution of  $\max\{d^T x \mid x \in P(q; k)\}$ , which proves that each vertex of  $P(q; k)$  lies in  $W(q; k)$ .

To complete the proof we must show that each point in  $W(q; k)$  is indeed a vertex of  $P(q; k)$ . To do this, assume that some  $w^s$  is a convex combination with *strictly positive* weights of other points  $z^m \in W(q; k)$ ,  $m = 1, \dots, N$ , i.e.  $w^s = \sum_{m=1}^N \lambda^m z^m$  with  $\lambda^m > 0$  and  $\sum_m \lambda^m = 1$ . Since all these points satisfy  $x_1 \leq q_1$  and  $w_1^s = q_1$ , we must have  $z_1^r = q_1$  for all  $r$ . Using a similar argument we can prove that  $z_j^r = q_j$  for  $j \leq s$ . All the points  $z^m$  lie in  $P(q; k)$  so they satisfy the valid inequality  $x(\mathbb{N}_k) \leq Q$ , and  $w^s$  satisfies this inequality with equality. This implies that  $z^r(\mathbb{N}_k) = Q$  and therefore by the concavity of the  $L$ -functions  $L_{z^r} \geq L_{w^s}$ . If one of these inequalities were strict, this would contradict that  $w^s = \sum_{m=1}^N \lambda^m z^m$ . Thus,  $L_{z^r} = L_{w^s}$  for all  $m$ , from which we get  $z^m = w^s$  for  $m = 1, \dots, N$ . This proves that  $w^s$  is an extreme point of  $P(q; k)$ , and it is therefore also a vertex of this polyhedron. Similarly, we can prove that each permutation of  $w^s$  is a vertex of  $P(q; k)$  and the proof is complete.  $\square$

An interesting consequence of Theorem 2 is a simple algorithm for solving  $\max \{c^T x \mid x \prec_k q\}$ . Initially the components of  $c$  are ordered decreasingly; this can be done in time  $O(n^2)$ . Assume for simplicity that  $c$  is nonincreasing. Next we determine an index  $m$  as in Lemma 1 and let  $s = m - 1$ . Then  $w^s$  is an optimal solution of the problem! Here the calculation of  $m$  is done in linear time by recursive calculations of tail averages. Thus the main work in the algorithm is simply the ordering of the objective function.

We remark that Theorem 2 can also be proved by a more direct method without using LP duality. In fact, one can show that if  $x$  is an optimal solution of (8) satisfying  $x(\mathbb{N}_s) = q(\mathbb{N}_s)$  for some  $s \in \{0, \dots, k-1\}$ , then we can find an optimal solution  $\bar{x}$  satisfying  $\bar{x}_j = q_j$  for  $j \leq s$ . By choosing  $s$  maximal with this property, one can find another optimal solution  $\hat{x}$  satisfying  $\hat{x}(\mathbb{N}_k) = \sum_{j=1}^k q_j$  and with all the remaining components  $\hat{x}_{s+1}, \dots, \hat{x}_n$  equal; this solution  $\hat{x}$  is precisely the  $q$ -average  $w^s$ .

From Theorem 2 we also obtain an inner description of weak  $k$ -majorization polyhedra by using the fact that the characteristic cone of  $P(q; k)$  is  $-\mathbb{R}^n$ .

**Corollary 3**  $P(q; k) = \text{conv}(W(q; k)) - \mathbb{R}^n$ , i.e.,  $x \prec_k q$  if and only if  $x \leq z$  for some  $z \in \mathbb{R}^n$  which is a convex combination of permuted  $q$ -

averages.

In Fig.1 we illustrate the intersection between  $P(q; k)$  and the nonnegative orthant for  $k = 2$ ,  $n = 3$  and  $q = (2, 1)$ . The different permuted  $q$ -averages are shown.

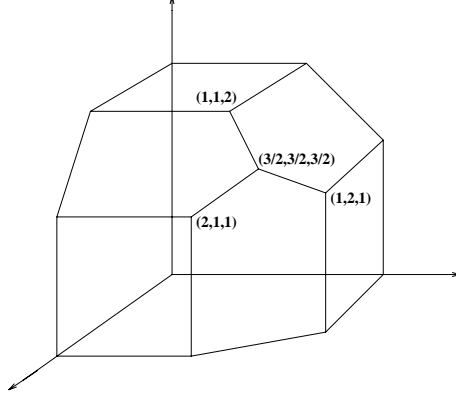


Figure 1: *Example,  $P(q; k) \cap \mathbb{R}^n$ .*

## 4 Weak $k$ -majorization and integrality

In this section we study integer linear programming problems with a weak  $k$ -majorization constraint. These problems may be motivated by applications concerning e.g. the distribution of indivisible “units” to locations or projects where it may be natural to impose a majorization constraint to assure a certain level of diversification. Our approach is to study the integer hull of the polyhedron  $P(q; k)$  and determine its vertices.

Hereafter  $k$  and  $n$  are integers satisfying  $k < n$  and the majorant  $q$  satisfies  $q_1 > \dots > q_k > 0$ . We are interested in the following integer linear programming problem

$$\max\{c^T x \mid x \prec_k q, x \text{ is integral}\}. \quad (10)$$

Note that the feasible set is here  $P(q; k) \cap \mathbb{Z}^n$ . If some  $c_j$  is negative, then (10) is unbounded, so we hereafter assume that  $c_1 \geq \dots \geq c_n$ . Note that the LP relaxation  $\max\{c^T x \mid x \in P(q; k)\}$  may have (a unique) optimal solution that is fractional. For instance, consider again the example  $n = 3$ ,  $k = 2$ ,  $q = (2, 1)$  and  $c^T = (1, 1, 1)$ . The optimal solution in  $P(q; k)$  is  $(3/2, 3/2, 3/2)$ , while the optimal integral solution is  $(2, 1, 1)$ .

We observe that the feasible set of (10) is unchanged if we perform integer round-down on each component of the majorant  $q$ . Thus we shall hereafter assume that  $q$  is integral (this also strengthens the LP relaxation).

The integer hull  $Q(q; k)$  of  $P(q; k)$  is the convex hull of the feasible set in (10), i.e.,

$$Q(q; k) = \text{conv}(\{x \in \mathbb{R}^n \mid x \prec_k q, x \text{ is integral}\}). \quad (11)$$

From general polyhedral theory (see [10]) we know that  $Q(q; k)$  is a polyhedron, i.e., it is the solution set of a finite set of linear inequalities.  $Q(q; k)$  is pointed and its characteristic cone is  $-\mathbb{R}^n$ . Furthermore it is easy to see that  $Q(q; k)$  is symmetric. Note that the problem (10) may be viewed as maximizing  $c^T x$  over  $Q(q; k)$ .

Some more notation is needed. Let  $m^{max} = \lfloor \bar{q}_{1:k} \rfloor = \lfloor Q/k \rfloor$  (recall that  $Q = \sum_{j=1}^k q_j$ ). For each nonnegative  $m \leq m^{max}$  we define the following magnitudes

- $s^m = \max\{s < k \mid \bar{q}_{s+1:k} \geq m\}$ ;
- $\Delta^m = \sum_{j=s^m+1}^k q_j - (k - s^m - 1)m$ ;
- $q^m = (q_1, \dots, q_{s^m}, \Delta^m, m, \dots, m) \in \mathbb{R}^n$ .

We call  $q^m$  a *rounded  $q$ -average* due to its near resemblance to the  $q$ -averages  $w^s$  (see section 3). The first  $s^m$  components are maximally spread out (w.r.t.  $q$ ) while the remaining components are minimally spread out: they are all equal except for the  $(s^m + 1)$ 'th which may be viewed as an “integrality correction”. Observe that if  $m \leq q_k$ , then  $s^m = k - 1$  and  $q^m = (q_1, \dots, q_k, m, \dots, m)$ . Finally, we note that if  $m := \bar{q}_{s+1:k}$  is integral, then  $s^m = s$  and  $q^m = w^s$ . Some basic properties of  $q^m$  are given next.

**Lemma 4** *Let  $m$  be a nonnegative integer with  $m \leq m^{max}$ . Then  $q^m$  is integral, nonincreasing and  $q^m \prec_k q$ . Furthermore  $q^m(\mathbb{N}_k) = Q$ .*

**Proof.** It is clear that  $q^m$  is integral. To prove that the components of  $q^m$  are nonincreasing we shall bound  $\Delta^m$ . Let  $s = s^m$ . Then we have  $\Delta^m = \sum_{j=s+1}^k q_j - (k - s - 1)m = (k - s)(\bar{q}_{s+1:k} - m) + m \geq m$  because  $k - s \geq 1$  and  $\bar{q}_{s+1:k} \geq m$  (due to the definition of  $s = s^m$ ). Furthermore  $\Delta^m = \sum_{j=s+1}^k q_j - (k - s - 1)m \leq \sum_{j=s+1}^k q_j - (k - s - 1)\bar{q}_{s+2:k} = \sum_{j=s+1}^k q_j - \sum_{j=s+2}^k q_j = q_{s+1} < q_s$ . Thus  $q_s > \Delta^m \geq m$  and  $q^m$  is nonincreasing. To prove the majorization, we first observe that when  $j \leq s$  we have  $q^m(\mathbb{N}_j) = q(\mathbb{N}_j)$  and by the first part of the proof  $q^m(\mathbb{N}_{s+1}) = q(\mathbb{N}_s) + \Delta^m < q(\mathbb{N}_s) + q_{s+1} = q(\mathbb{N}_{s+1})$ . This shows that  $L_{q^m} \leq L_q$  on  $[0, (s+1)k]$ . Next,  $q^m(\mathbb{N}_k) = q(\mathbb{N}_s) + \Delta^m + (k - s - 1)m = q(\mathbb{N}_s) + \sum_{j=s+1}^k q_j = q(\mathbb{N}_k) = Q$ , so  $L_{q^m}(k) = L_q(k)$ . This combined with the concavity of  $L_q$  and the fact that  $L_{q^m}$  is linear on  $[(s+1)/k, 1]$  implies that  $L_q \geq L_{q^m}$  on  $[(s+1)/k, 1]$ . Thus  $L_q \geq L_{q^m}$  on  $[0, 1]$  and therefore  $q^m \prec_k q$ .  $\square$

Note that  $q^{m_0} \geq q^m$  for  $m_0 := q_k > m$ , so we have  $c^T q^{m_0} \geq c^T q^m$  as the objective function  $c$  is nonnegative. We shall therefore hereafter restrict the attention to  $m \in M := \{q_k, \dots, m^{max}\}$ . Since  $q$  is strictly decreasing, the tail averages are ordered  $\bar{q}_{k:k} < \bar{q}_{k-1:k} < \dots < \bar{q}_{1:k}$ . Therefore the closed intervals  $I_s := [\bar{q}_{s+1:k}, \bar{q}_{s:k}]$  for  $s = 1, \dots, k-1$  are all

nonempty and consecutive intervals intersect in one point, namely a tail average of  $q$ . We say that  $m \in M$  is  $q$ -extreme if  $m$  is either the smallest or largest integer in one of the intervals  $I_s$ ,  $s = 1, \dots, k-1$ . Some properties of the  $q$ -extreme integers are collected in the next lemma; they are all easy consequences of the definition of  $s^m$ . We let  $M^{ex}$  denote the set of  $q$ -extreme integers.

**Lemma 5** *Let  $m \in M$ . Then we have*

- (i)  $m \in I_{s^m+1}$ ;
- (ii) if  $m \in M^{ex}$ , then  $m$  equals either  $\lfloor \bar{q}_{s^m+1:k} \rfloor$  or  $\lceil \bar{q}_{s^m+2:k} \rceil$ ;
- (iii)  $m \in M^{ex}$  if and only if  $m$ 's is obtained from some tail average  $\bar{q}_{s:k}$  by integer rounding either up or down.

We are now in a position to give the first main result of this section.

**Proposition 6** *Each vertex of  $Q(q; k)$  is a permutation of some rounded  $q$ -average  $q^m$  with  $m$  being  $q$ -extreme.*

**Proof.** We consider the ILP problem (10) and may assume that  $c_1 \geq \dots \geq c_n \geq 0$  (otherwise a suitable permutation would be introduced below). We shall prove that some rounded  $q$ -average is an optimal solution of this problem. From the rearrangement inequality (1) it follows that there exists an optimal solution  $x$  satisfying  $x_1 \geq \dots \geq x_n \geq 0$ , so we may consider the ILP problem

$$\max\{c^T x \mid x \prec_k q, x_1 \geq \dots \geq x_n \geq 0, x \text{ is integral}\}. \quad (12)$$

Note that when  $x$  is nonincreasing,  $x \prec_k q$  holds if and only if  $x(\mathbb{N}_r) \leq q(\mathbb{N}_r)$  for  $r = 1, \dots, k$ . Thus the variables  $x_{k+1}, \dots, x_n$  do not enter any of these constraints and we may assume that  $x_k = \dots = x_n$  (as  $c \geq 0$ ). For each  $m \in M$ , we consider the problem obtained from (12) by adding the constraints  $x_k = \dots = x_n = m$ . We shall solve this problem directly using previous results, and first we note that the problem may be considered as one in  $k$  variables:

$$\max\{d^T x \mid x \prec_k q, x_1 \geq \dots \geq x_k = m, x \text{ is integral}\}. \quad (13)$$

where the objective function  $d$  is given by  $d_j = c_j$  for  $j \leq k-1$  and  $d_k = \sum_{j=k}^n c_j$ .

We claim that there exists an optimal solution  $x$  of (13) satisfying  $x(\mathbb{N}_k) = q(\mathbb{N}_k)$ . To prove this, consider an optimal solution with  $x(\mathbb{N}_k)$  largest possible. Assume that  $x(\mathbb{N}_k) < q(\mathbb{N}_k)$ ; we shall deduce a contradiction. Observe that  $x(\mathbb{N}_k) \leq q(\mathbb{N}_k) - 1$  as  $x$  is integral. If  $x_s > x_{s+1} = \dots = x_k$  for some  $s < k$ , we let  $\tilde{x}$  be the solution obtained from  $x$  by increasing the  $(s+1)$ 'th component by 1. Clearly  $\tilde{x}$  is decreasing. Furthermore, as  $q$  is strictly decreasing and  $x_{s+1} = \dots = x_k$ , the  $L$ -curves  $L_x$  and  $L_q$  can only intersect on the interval  $[s/k, 1]$  in one point, namely for  $j = s/k$ . Thus  $L_x$  is strictly below  $L_q$  in  $[(s+1)/k, 1]$ , and therefore  $L_{\tilde{x}} \leq L_q$  in this interval. Also, in  $[0, s/k]$   $L_{\tilde{x}} = L_x$ , so  $\tilde{x} \prec_k q$  which implies that  $\tilde{x}$  is also optimal. In addition  $\tilde{x}(\mathbb{N}_k) = x(\mathbb{N}_k) + 1$ , which contradicts the maximality of  $x(\mathbb{N}_k)$ . It remains to discuss the case when

$x_1 = \dots = x_k$ . As  $q$  is strictly decreasing, we must have  $x(\mathbb{N}_j) < q(\mathbb{N}_j)$  for all  $j$ . Consequently, if we define the new solution  $\tilde{x}$  from  $x$  by increasing the first variable by 1, we obtain, as in the first case discussed, a new optimal solution with  $\tilde{x}(\mathbb{N}_k) = x(\mathbb{N}_k) + 1$  again contradicting the maximality of  $x(\mathbb{N}_k)$ . This proves the claim.

Consider (as in the claim above) an optimal solution  $x \in \mathbb{R}^k$  of (13) with  $x(\mathbb{N}_k) = q(\mathbb{N}_k)$ . The feasibility of  $x$  implies that  $L_x \leq L_q$  and  $L_x \leq L'$  on  $[0, 1]$  where  $L'(\theta) := Q - km(1 - \theta)$  for  $\theta \in [0, 1]$ . Therefore  $L \leq L'' := \min\{L_q, L'\}$ . It is easy to check that  $L'' = L_{q^m}$ , so we get  $x \prec_k P_k(q^m)$  where  $P_k(q^m)$  is the vector (projection) containing the first  $k$  components of  $q^m$ . (Remark:  $L_q$  is a function on  $[0, 1]$ , although  $q^m \in \mathbb{R}^n$ ). From the discussion in section 2 on optimizing over  $P(q; n)$  using the greedy algorithm, it is clear that  $P_k(q^m)$  maximizes  $d^T z$  for  $z \in \mathbb{R}^k$  satisfying  $z \prec_k P_k(q^m)$ . Thus  $d^T P_k(q^m) \geq d^T x$ , and the optimality of  $x$  in problem (13) gives that  $P_k(q^m)$  is also an optimal solution to that problem.

Finally, by taking the maximum optimal value of the problems (13) over all  $m$  we solve the original problem (10). The optimal value of (13) is (we let  $s = s^m$ )  $\sum_{j=1}^k c_j P_k(q^m)_j = \sum_{j=1}^s c_j q_j + c_{s+1} \{\sum_{j=s+1}^k q_j - (k - s - 1)m\} + \sum_{j=s+2}^n c_j m = \sum_{j=1}^s c_j q_j + c_{s+1} \sum_{j=s+1}^k q_j + m \{\sum_{j=s+2}^n c_j - (k - s - 1)c_{s+1}\} = \alpha(s) + m\beta(s)$ , where  $\alpha(s)$  and  $\beta(s)$  are defined in the obvious way. Observe that  $\alpha$  and  $\beta$  depends on  $m$  but only through  $s = s^m$ . Thus the optimal value for those  $m$  with  $s^m = s'$  (for some  $s'$ ) is achieved by choosing  $m$  either maximal (if  $\beta \geq 0$ ) or minimal (if  $\beta < 0$ ) with  $s^m = s'$ , i.e.  $m$  is extreme. Based on the initial discussions in this proof, we can conclude that for some  $q$ -extreme  $m \in M$  the rounded  $q$ -average  $q^m$  is an optimal solution of (10) which proves that each vertex of  $Q(q; k)$  must be a rounded  $q$ -average as desired.  $\square$

The previous proposition shows the form of each vertex, but it remains to decide if *all* the solutions  $q^m$  with  $m$  extreme are vertices of  $Q(q; k)$ . The answer turns out to be affirmative, and to prove this we shall first give an optimization result which is useful later as well. We study the problem  $\max\{c^T x \mid x \in Q(q; k)\}$  for a special stair-case objective function. For each real number  $\gamma \geq 1$  and each pair of integers  $s, t$  with  $0 \leq s < k < t \leq n$ , we define  $c^{s,t}(\gamma) \in \mathbb{R}^n$  by

$$c^{s,t}(\gamma) = \begin{cases} \gamma & \text{for } j = 1, \dots, s, \\ 1 & \text{for } j = s+1, \dots, t, \\ 0 & \text{for } j = t+1, \dots, n, \end{cases} \quad (14)$$

and we consider the problem

$$\max\{(c^{s,t}(\gamma))^T x \mid x \in Q(q; k)\}. \quad (15)$$

**Lemma 7** *Let  $s$  and  $t$  be as above. Define  $\bar{m} = \lfloor \bar{q}_{s+1:k} \rfloor$ ,  $\hat{m} = \bar{m} + 1$ ,  $\mu = (k - s)\hat{m} - \sum_{j=s+1}^k q_j$ ,  $\gamma_1 = (t - s)/(k - s)$  and  $\gamma_2 = 1 + (t - k)/\mu$ . Finally let  $M^*(\gamma)$  denote the set of  $m \in M$  such that  $q^m$  is an optimal solution of (15).*

Then  $\mu \geq 1$  and  $\gamma_2 \geq \gamma_1 > 1$ ; in fact  $\gamma_2 > \gamma_1$  holds if and only if  $\bar{q}_{s+1:k}$  is fractional. Furthermore  $M^*(\gamma)$  is given by the following expressions:

- (i) If  $1 \leq \gamma < \gamma_1$ , then  $M^*(\gamma) = \{m^{max}\}$ .
- (ii) If  $\gamma = \gamma_1$ , then  $M^*(\gamma) = \{\hat{m}, \dots, m^{max}\}$ .
- (iii) If  $\gamma_1 < \gamma < \gamma_2$ , then  $M^*(\gamma) = \{\hat{m}\}$ .
- (iv) If  $\gamma = \gamma_2$ , then  $M^*(\gamma) = \{\bar{m}, \hat{m}\}$ .
- (v) If  $\gamma > \gamma_2$ , then  $M^*(\gamma) = \{\bar{m}\}$ .

**Proof.** Let  $c = c^{s,t}(\gamma)$ . From the definition of  $\bar{m}$  we get  $\bar{m} + 1 > \bar{q}_{s+1:k} = (1/(k-s)) \sum_{j=s+1}^k q_j$  and thus  $\mu = (k-s)(\bar{m} + 1) - \sum_{j=s+1}^k q_j > 0$ . As  $\mu$  is integral, we get  $\mu \geq 1$ . Note that  $\gamma_1$  is well-defined as  $s < k$  and also  $\gamma_1 = (t-s)/(k-s) > 1$  because  $t > k$ . Furthermore, we have  $\bar{m} + 1 \leq \bar{q}_{s+1:k} + 1$ , which implies  $\mu \leq k-s$  and the inequality is strict if  $\bar{q}_{s+1:k}$  is fractional. Thus  $\gamma_2 = 1 + (t-k)/\mu \geq 1 + (t-k)/(k-s) = (t-s)/(k-s) = \gamma_1$  and here we see that  $\gamma_2 > \gamma_1$  if  $\bar{q}_{s+1:k}$  is fractional.

Consider the optimization problem (15). From Proposition 6, we know that there is an optimal solution of the form  $q^m$  for some  $m \in M^{ex}$ . We have

$$c^T q^m = (\gamma - 1) \sum_{j=1}^s q_j^m + Q + (t-k)m. \quad (16)$$

as  $q_j^m = m$  for  $j \geq k$  and  $q^m(\mathbb{N}_k) = Q$ . This implies that the following set of inequalities hold

$$c^T q^{q_k} < \dots < c^T q^{\bar{m}}. \quad (17)$$

In fact, for  $m \leq \bar{m}$  we get  $s^m \geq s$  and therefore  $q_j^m = q_j$  for  $j = 1, \dots, s$ . Thus from (16) we get  $c^T q^m = (\gamma - 1) \sum_{j=1}^s q_j + (t-k)m + Q$  which is strictly increasing as a function of  $m$  (due to  $t > k$ ), and all the inequalities in (17) follow.

Next, let  $m > \bar{m}$ , so  $s^m < s$ . Then  $q_{s+1}^m = \dots = q_k^m = m$  and  $\sum_{j=1}^s q_j^m = \sum_{j=1}^k q_j^m - \sum_{j=s+1}^k q_j^m = Q - (k-s)m$ . By inserting this in (16) we get  $(\gamma - 1)(Q - (k-s)m) + (t-k)m + Q = \psi m + \gamma Q$  where  $\psi = t-k - (\gamma - 1)(k-s)$ . Therefore  $\psi$  is positive (resp. zero, negative) iff  $\gamma$  is smaller than (resp. equal to, larger than)  $\gamma_1 = (t-s)/(k-s)$ . Thus if  $1 \leq \gamma < \gamma_1$ , then  $\psi > 0$  and therefore  $c^T q^{\hat{m}} < \dots < c^T q^{m^{max}}$ . Next, if  $\gamma = \gamma_1$ , then  $\psi = 0$  and  $c^T q^{\hat{m}} = \dots = c^T q^{m^{max}}$ . And, finally, if  $\gamma > \gamma_1$ , then  $\psi < 0$  and  $c^T q^{\bar{m}} > \dots > c^T q^{m^{max}}$ .

It remains to compare  $q^{\bar{m}}$  and  $q^{\hat{m}}$ . From (16) we obtain after some calculations that  $c^T q^{\bar{m}} - c^T q^{\hat{m}} = (\gamma - 1)\mu - (t-k)$ . Thus (i) if  $\gamma > 1 + (t-k)/\mu = \gamma_2$ , then  $c^T q^{\bar{m}} > c^T q^{\hat{m}}$ , (ii) if  $\gamma = \gamma_2$ , then  $c^T q^{\bar{m}} = c^T q^{\hat{m}}$ ; and (iii) if  $\gamma < \gamma_2$ , then  $c^T q^{\bar{m}} < c^T q^{\hat{m}}$ .

The lemma now follows by putting all these inequalities together for varying  $\gamma$ . □

The complete characterization of the vertices of  $Q(q; k)$  can now be given.

**Theorem 8** *The vertex set of  $Q(q; k)$  consists of the vectors that can be obtained as permutations of some  $q^m$  with  $m$  being  $q$ -extreme.*

**Proof.** On the basis of Proposition 6 we only need to prove that when  $m$  is  $q$ -extreme the point  $q^m$  is a vertex of  $Q(q; k)$ , because then each permutation of  $q^m$  is also a vertex (due to the symmetry of  $Q(q; k)$ ). It therefore suffices to prove that if  $m \in M^{ex}$  then  $q^m$  is the *unique* optimal solution of some ILP problem (10).

Let  $m$  be  $q$ -extreme, and recall the form of the  $q$ -extreme integers given in Lemma 5. Assume first that  $m$  equals some tail average  $\bar{q}_{s+1:k}$  of  $q$  (so this average must be integral). Then  $q^m = w^s$  and it is therefore by Theorem 2 a vertex of  $P(q; k)$ . But since  $Q(q; k) \subseteq P(q; k)$ , it follows that  $q^m$  is also a vertex of  $Q(q; k)$ . Next, assume that  $m = \lfloor \bar{q}_{s+1:k} \rfloor$  where  $\bar{q}_{s+1:k}$  is fractional. The remaining case where  $m = \lceil \bar{q}_{s+1:k} \rceil$  is treated similarly.

Choose an objective function  $c \in \mathbb{R}^n$  satisfying (i)  $c_r > \sum_{j=r+1}^n c_j$  for  $r = 1, \dots, s$ , (ii)  $c_{s+1} = \gamma$  where  $1 < \gamma < (n - s - 1)/(k - s - 1)$  and (iii)  $c_{s+2} = \dots = c_n = 1$ . Let  $x$  be an optimal solution of problem (10). We claim that  $x$  must satisfy  $x_j = q_j$  for  $j \leq s$ . In fact, if  $x_1 < q_1$ , we find another feasible solution in (10) by increasing  $x_1$  by a “small” amount  $\epsilon$  and reducing all other components by  $\epsilon$ . But the objective function value for this new solution is strictly higher than  $c^T x$  as  $c_1 > \sum_{j=2}^n c_j$ , which contradicts the optimality of  $x$ . This proves that  $x_1 = q_1$ , and with similar arguments the claim follows.

In order to determine the remaining components  $x_{s+1}, \dots, x_n$  of  $x$ , we observe that these must be an optimal solution of the “reduced” problem  $\max \{(c')^T x' \mid x' \prec_{k-s} q', x' \text{ integral}\}$  where  $x' = (x_{s+1}, \dots, x_n)$ ,  $c' = (c_{s+1}, \dots, c_n)$  and  $q' = (q_{s+1}, \dots, q_k)$ . But this problem can be solved using Lemma 7 where the parameters  $n$ ,  $k$ ,  $s$  and  $t$  of that lemma are set to  $n - s$  (the number of variables),  $k - s$ , 1 and  $n - s$ , respectively and the objective function is  $c' = (\gamma, 1, \dots, 1)$ . With this choice we get  $\gamma_1 = (n - s - 1)/(k - s - 1)$  so  $1 < \gamma < \gamma_1$ , and from Lemma 7(i) we have  $M^*(\gamma) = \{m^{max}\} = \{\lfloor \bar{q}_{s+1:k} \rfloor\}$ . As  $m = \lfloor \bar{q}_{s+1:k} \rfloor$ , it follows that  $q^m$  is an optimal solution of (10) and that no other  $q$ -average can be optimal, i.e.,  $c^T q^{m'} < c^T q^m$  for  $m' \neq m$ . Therefore, if  $z$  is another optimal solution, then  $z$  must be a permutation of  $q^m$  (confer Proposition 6). But since  $c_{s+1} > c_j = 1$  and  $q_{s+1}^m > q_j^m = m$  for all  $j > s + 1$ , we derive that  $z = q^m$ . Thus  $q^m$  is the unique optimal solution of (10) and it is therefore a vertex of  $Q(q; k)$  which completes the proof.  $\square$

**Examples.** Let  $k = 3$ ,  $n = 5$  and  $q = (7, 2, 1)$ . Then the rounded  $q$ -averages are  $q^1 = (7, 2, 1, 1, 1)$ ,  $q^2 = (6, 2, 2, 2, 2)$  and  $q^3 = (4, 3, 3, 3, 3)$ , and the vertices of  $Q(q; k)$  are all permutations of these points. As another example let  $k = 4$ ,  $n = 6$  and  $q = (19, 12, 5, 3)$ . Then the tail averages of  $q$  are 3, 4,  $20/3$  and  $39/4$  and the  $q$  extreme integers are therefore  $M^{ex} = \{3, 4, 6, 7, 9\}$ . The  $q$ -averages are  $q^3 = (19, 12, 5, 3, 3, 3)$ ,  $q^4 = (19, 12, 4, 4, 4, 4)$ ,  $q^6 = (19, 8, 6, 6, 6, 6)$ ,  $q^7 = (18, 7, 7, 7, 7, 7)$  and  $q^9 = (12, 9, 9, 9, 9, 9)$ .

How many vertices does  $Q(q; k)$  have, or equivalently, how many different rounded  $q$ -averages with  $m$   $q$ -extreme can we have? (The first number



is obtained from the latter by also counting permutations). We give simple bounds on this number based on the increments of the components of  $q$ .

**Proposition 9**  $\lceil \frac{k}{2} \rceil \leq |M^{ex}| \leq 2k$ .

**Proof.** The upper bound is trivial, so we only need to prove the lower bound. A certain property of tail averages is useful. Assume that  $q_s \geq q_{s+1} + \epsilon$  for  $s = 1, \dots, k-1$ . Then, for  $s = 1, \dots, k-1$ , we have

$$\bar{q}_{s:k} \geq \bar{q}_{s+1:k} + \frac{\epsilon}{2} \quad (18)$$

In fact, it follows from the assumption that  $q_s - q_j \geq (j-s)\epsilon$  for  $j \geq s$  and therefore  $q_s - \bar{q}_{s:k} = (1/(k-s+1)) \sum_{j=s}^k (q_s - q_j) \geq (k-s)\epsilon/2$ . Applying this inequality and that  $q_s \geq q_{s+1} + \epsilon$  we obtain  $\bar{q}_{s:k} - \bar{q}_{s+1:k} = q_s/(k-s+1) + \bar{q}_{s+1:k}(k-s)/(k-s+1) - \bar{q}_{s+1:k} = (q_s - \bar{q}_{s+1:k})/(k-s+1) \geq (\epsilon + q_{s+1} - \bar{q}_{s+1:k})/(k-s+1) \geq (\epsilon + \epsilon(k-s+1)/2)/(k-s+1) = \epsilon/2$  which proves (18). Since  $q$  is assumed to be strictly decreasing, (18) holds with  $\epsilon = 1$ , and applying that inequality twice gives  $\bar{q}_{s:k} \geq \bar{q}_{s+2:k} + 1$ . Thus there must be an integer strictly between these two tail averages or they are both integers. This proves that  $\lceil \bar{q}_{s:k} \rceil < \bar{q}_{s-2:k}$  and  $\lfloor \bar{q}_{s:k} \rfloor > \bar{q}_{s+2:k}$  and therefore  $M^{ex} \cap (I_s \cup I_{s+1})$  is nonempty. But then the lower bound in the proposition follows as in the “worst” case every second  $I_s$  is empty.  $\square$

Consider again the ILP problem (10), or equivalently the LP problem  $\max \{c^T x \mid x \in Q(q; k)\}$ . Since we know all the vertices of  $Q(q; k)$  and due to their “nice” structure, we can solve this problem as follows. First we order the components of  $c$  nonincreasingly. Assume for simplicity that  $c_1 \geq \dots \geq c_n$ , and by the rearrangement inequality there exists an optimal solution  $x$  which is nonincreasing. In fact, one of the vertices  $q^m$  is optimal, and we could simply calculate  $c^T q^m$  for each  $q$ -extreme  $m$  and thereby solve the problem. Note that as Proposition 9 says there are at most  $k$  vertices we need to compare, this algorithm is polynomial. In the next section we shall do even better by giving a rather similar algorithm which only compares *two* of these vertices. This is done in connection with a study of the facets of  $Q(q; k)$ .

## 5 Linear description of the integer hull of weak $k$ -majorization polyhedra

In this section we study the facets (maximal faces) of the polyhedron  $Q(q; k)$  defined in (11). The final goal is to determine a complete and nonredundant linear description of this polyhedron. Initially we study simple facets known already from the polyhedron  $P(q; k)$ , and next we study additional facets that are due to the integrality of  $Q(q; k)$ . Again we assume that the integers  $n$  and  $k$  satisfy  $k < n$  and that the majorant  $q$  is an integral vector with strictly decreasing components.

## 5.1 Basic facets

We describe basic facets of  $Q(q; k)$  corresponding to inequalities defining  $P(q; k)$ .

**Proposition 10** *The polyhedron  $Q(q; k)$  is full-dimensional and for each  $S \subset \mathbb{N}_n$  with  $r = |S| \leq k$  the inequality*

$$x(S) \leq q(\mathbb{N}_r) \quad (19)$$

*defines a facet of  $Q(q; k)$ .*

**Proof.** The full-dimensionality of  $Q(q; k)$  follows from the fact that the  $n+1$  vectors  $e_j$ ,  $j \in \mathbb{N}_n$  and 0 all lie in  $Q(q; k)$  and are affinely independent. Let  $S \subset \mathbb{N}_n$  with  $r := |S| \leq k$  and let  $F$  be the face of  $Q(q; k)$  induced by the valid inequality  $x(S) \leq q(\mathbb{N}_r)$ , i.e.,  $F = \{x \in Q(q; k) \mid x(S) = q(\mathbb{N}_r)\}$ . Then  $F$  contains all permutations  $x$  of  $q$  such that the  $r$  (largest) components  $q_1, \dots, q_r$  of  $q$  are mapped to  $S$  (that is:  $(x_i : i \in S)$  is a permutation of  $(q_1, \dots, q_r)$ ). Let  $b^T x \leq \beta$  be a valid inequality for  $Q(q; k)$  that induces a facet  $F(b) := \{x \in Q(q; k) \mid b^T x = \beta\}$  with  $F \subseteq F(b)$ . By comparing adjacent permutations (like  $(q_1, q_2, q_3, \dots, q_n)$  and  $(q_2, q_1, q_3, \dots, q_n)$ ), it follows that  $b_i = b_j$  for all  $i, j \in S$  (here the details are as in the proof of Proposition 13). So, for some constant  $d$ , we have  $b_j = d$  for all  $j \in S$ . Next, consider the vector  $q'$  obtained from  $q$  by setting the  $j$ 'th component to 0 for some  $j \in \mathbb{N}_n \setminus S$ . Then both  $q$  and  $q'$  lie in  $F$  and therefore also in  $F(b)$ , and this implies that  $b_j = 0$ ,  $j \in \mathbb{N}_n \setminus S$ . Thus  $b^T x \leq \beta$  has the form  $dx(S) \leq \beta$ . Since  $q \in F$ , we obtain  $\beta = dq(\mathbb{N}_r)$ . This proves that  $b^T x \leq \beta$  is positive scalar multiple of  $x(S) \leq q(\mathbb{N}_r)$ , and therefore  $x(S) \leq q(\mathbb{N}_r)$  defines a facet of  $Q(q; k)$ .  $\square$

We call each inequality in (19) a *set size inequality*.

We next study whenever the set size inequalities give a complete linear description of  $Q(q; k)$ , or equivalently, whenever  $Q(q; k)$  and  $P(q; k)$  coincide.

From Proposition 6, we know that all vertices of  $P(q; k)$  are permutations of the points  $w^s = (q_1, \dots, q_s, \bar{q}_{s+1:k}, \dots, \bar{q}_{s+1:k})$ , and this vector is integral if and only if  $k - s \mid \sum_{j=s+1}^k q_j$ . Note that  $\bar{q}_{k:k} = q_k$  is always integral since  $q$  is integral. It follows that  $P(q; k)$  is an integral polyhedron, i.e.  $P(q; k) = Q(q; k)$ , if and only if  $k - s \mid \sum_{j=s+1}^k q_j$  for  $s = 0, \dots, k - 2$ . From this it is easy to derive the following characterization.

**Proposition 11** *The polyhedron  $P(q; k)$  is integral if and only if  $q$  can be obtained by the following procedure: let  $q_k$  be any positive integer and for  $r = k - 1, \dots, 1$ , choose  $q_r > q_{r+1}$  such that  $q_r = m_r(k - r + 1) - \sum_{j=r+1}^k q_j$  for some positive integer  $m_r$ .*

Whenever  $q$  is not constructed as in Proposition 11, the polyhedron  $P(q; k)$  has some fractional vertex, and therefore  $Q(q; k)$  is strictly contained in  $P(q; k)$ . The remaining part of the paper is devoted to the question of finding additional inequalities that are needed to get a complete linear

description of  $Q(q; k)$  in the general situation. From Motzkin's decomposition theorem for polyhedra (see [10]) the *existence* of such a linear system is clear.

## 5.2 The $q$ -average inequalities

Let  $s$  and  $t$  be integers satisfying  $0 \leq s < k < t \leq n$ . Define  $\delta^s = \sum_{j=s+1}^k q_j - (k-s)\lfloor \bar{q}_{s+1:k} \rfloor$  which is the rest modulo  $k-s$  of  $\sum_{j=s+1}^k q_j$ . Consider the inequality

$$(t-s-\delta^s) \sum_{j=1}^s x_j + (k-s-\delta^s) \sum_{j=s+1}^t x_j \leq \alpha^{s,t} \quad (20)$$

where  $\alpha^{s,t} = (t-k) \sum_{j=1}^s q_j + (k-s-\delta^s)(Q + (t-k)\lfloor \bar{q}_{s+1:k} \rfloor)$ . We remark that  $\alpha^{s,t}$  equals the value we obtain by inserting  $x = q^{\bar{m}}$  where  $\bar{m} = \lfloor \bar{q}_{s+1:k} \rfloor$  on the left-hand-side of (20). We call an inequality of the form (20) a  *$q$ -average inequality*, and it may be written as  $(a^{s,t})^T x \leq \alpha^{s,t}$ , where we define

$$a^{s,t} = \begin{cases} t-s-\delta^s & \text{for } j = 1, \dots, s, \\ k-s-\delta^s & \text{for } j = s+1, \dots, t, \\ 0 & \text{for } j = t+1, \dots, n. \end{cases} \quad (21)$$

Note here that  $a_s^{s,t} > a_{s+1}^{s,t} > 0$  as  $t > k$  and  $\delta^s < k-s$  (see the definition of  $\delta^s$ ).

**Lemma 12** *The  $q$ -average inequality (20) is valid for  $Q(q; k)$ .*

**Proof.** It follows from Lemma 7 (using the notation of that lemma) that

$$(c^{st})^T x \leq (c^{st})^T q^{\bar{m}} \quad (22)$$

where  $\bar{m} = \lfloor \bar{q}_{s+1:k} \rfloor$  is a valid inequality for  $Q(q; k)$ . We have  $\mu = (k-s)\bar{m} - \sum_{j=s+1}^k q_j = k-s + (k-s)\lfloor \bar{q}_{s+1:k} \rfloor - \sum_{j=s+1}^k q_j = k-s-\delta^s$  and  $\mu\gamma_2 = \mu + t - k = t-s-\delta^s$ , which proves that  $a^{s,t} = \mu c^{st}(\gamma_2)$  and  $\mu \geq 1$ . Therefore (20) is a positive scalar multiple of (22) and consequently the  $q$ -average inequality (20) is valid for  $Q(q; k)$ .  $\square$

Let  $\pi$  be a permutation on  $\mathbb{N}_n$  and recall the notation  $\pi x = (x_{\pi(1)}, \dots, x_{\pi(n)})$  for  $x \in \mathbb{R}^n$ . Since the polyhedron  $Q(q; k)$  is symmetric, it is clear that the permuted  $q$ -average inequality  $(\pi a^{s,t})^T x \leq \alpha^{s,t}$ , is also valid for  $Q(q; k)$ .

Consider the special case of the  $q$ -average inequalities obtained by setting  $s = 0$ ; this leads to the inequality  $(k-\delta^0) \sum_{j=1}^t x_j \leq (k-\delta^0)(Q + (t-k)\lfloor \bar{q}_{1:k} \rfloor)$  or equivalently

$$x(\mathbb{N}_t) \leq Q + (t-k)\lfloor \bar{q}_{1:k} \rfloor \quad (23)$$

We call each such inequality an *extended set size inequality* since it “extends” the set size inequalities to sets of cardinality larger than  $k$ .

**Example.** Consider again the example of section 4 where  $k = 3$ ,  $n = 5$  and  $q = (7, 2, 1)$ . We get the extended set size inequalities  $x(\mathbb{N}_4) \leq 13$  and  $x(\mathbb{N}_5) \leq 16$ . Both these inequalities cut off the fractional  $q$ -average  $w^0 = (10/3, \dots, 10/3)$ . Other  $q$ -average inequalities are  $2x_1 + x_2 + x_3 + x_4 \leq 18$  (obtained for  $s = 1$ ,  $t = 4$ ) and  $3x_1 + x_2 + x_3 + x_4 + x_5 \leq 26$  (for  $s = 1$ ,  $t = 5$ ).

We next characterize which of  $q$ -average inequalities that are nonredundant.

**Proposition 13** *The  $q$ -average inequality  $(\pi a^{s,t})^T x \leq \alpha^{s,t}$  defines a facet of  $Q(q; k)$  if and only if  $\bar{q}_{s+1:k}$  is fractional, i.e.  $k - s$  does not divide  $\sum_{j=s+1}^k q_j$ .*

**Proof.** It follows from the symmetry of  $Q(q; k)$  that it suffices to prove the result when  $\pi$  is the identity transformation.

Assume first that  $\bar{q}_{s+1:k}$  is fractional, and define  $\bar{m} = \lfloor \bar{q}_{s+1:k} \rfloor$  and  $\hat{m} = \bar{m} + 1$ . Then  $\Delta^{\bar{m}} > \bar{m}$ . Let  $F := \{x \in Q(q; k) \mid (a^{s,t})^T x = \alpha^{s,t}\}$  be the face of  $Q(q; k)$  induced by the (valid)  $q$ -average inequality. There exists a facet defining inequality  $b^T x \leq \beta$  such that  $F \subseteq F(b) = \{x \in Q(q; k) \mid b^T x = \beta\}$ . We shall prove that  $F = F(b)$ . Note first that we must have  $b \geq 0$  since the characteristic cone of  $Q(q; k)$  is  $-\mathbb{R}^n$ . From the proof of Lemma 12, it follows that both the rounded  $q$ -averages  $\bar{x} := q^{\bar{m}}$  and  $\hat{x} := q^{\hat{m}}$  lie in the face  $F$  (as  $M(\gamma_2) = \{\bar{m}, \hat{m}\}$ ; see Lemma 7). Therefore,  $\bar{x}, \hat{x} \in F(b)$ . Let  $i, j \leq s$  with  $i \neq j$  and let  $\tilde{x}$  be the vector obtained from  $\bar{x}$  by permuting the  $i$ 'th and  $j$ 'th components. Since  $a_i^{s,t} = a_j^{s,t}$ , it follows that  $\tilde{x} \in F$  and therefore  $\tilde{x} \in F(b)$ . So  $0 = \beta - b^T \bar{x} = b^T \tilde{x} - b^T \bar{x} = b_i(q_i - q_j) + b_j(q_j - q_i) = (b_i - b_j)(q_i - q_j)$ , so (as  $q_i > q_j$ ) we get  $b_i = b_j$  which then holds for all  $i, j \leq s$ . A similar argument can be used to prove that  $b_i = b_j$  for  $i, j \in \{s+1, \dots, t\}$  by using the fact that  $\bar{x}_{s+1} > \bar{x}_{s+2}$ , i.e. that  $\Delta^{\bar{m}} > \bar{m}$  (due to the fractionality of  $\bar{q}_{s+1:k}$ ). Furthermore, we must have  $b_i = 0$  for  $i > t$  because  $a_i^{s,t} = 0$  and  $\text{char.cone}(Q(q; k)) = -\mathbb{R}^n$ .

Summing up, there are constants  $b'$  and  $b''$  such that  $b$  satisfies  $b_i = b'$  for  $i \leq s$ ,  $b_i = b''$  for  $i = s+1, \dots, t$  and  $b_i = 0$  for  $i > t$ . It remains to relate  $b'$  and  $b''$ . Since  $\bar{x}, \hat{x} \in F \subseteq F(b)$ , we have  $b^T \bar{x} = b^T \hat{x}$  which implies  $b'(k - s - \delta^s) = b''(t - s - \delta^s)$ . Note that  $b' \neq 0$ ; otherwise we would get  $b = 0$  contradicting that  $b^T x \leq \beta$  defines a facet of  $Q(q; k)$ . Thus  $b'/b'' = (t - s - \delta^s)/(k - s - \delta^s)$  which proves that  $b = \lambda a^{s,t}$  for some positive scalar  $\lambda$  (recall that  $b \geq 0$ ). Therefore  $F = F(b)$  and  $(a^{s,t})^T x \leq \alpha^{s,t}$  is nonredundant.

Next, assume that  $\bar{m} := \bar{q}_{s+1:k}$  is integral and we shall prove that (20) is redundant. We get  $\delta^s = 0$  and (20) becomes

$$(t-s) \sum_{j=1}^s x_j + (k-s) \sum_{j=s+1}^t x_j \leq (t-s)Q \quad (24)$$

as  $\alpha^{s,t} = (t-s) \sum_{j=1}^s q_j + (k-s) \sum_{j=s+1}^t \bar{m} = (t-s)Q$ . Let  $C \in \mathbb{R}^{t-s, t-s}$  be the 0/1-circulant matrix  $C_{k-s, t-s}$  (see the notation in section 1) and let  $E \in \mathbb{R}^{t-s, s}$  be an all 1 matrix. Finally, let  $b \in \mathbb{R}^{t-s}$  be a vector

with all elements being  $Q$  and  $D \in \mathbb{R}^{t-s, n-t}$  be a all 0 matrix. Then the linear system  $[E, C, D]x \leq b$  consists of  $t-s$  set size inequalities (size  $k$ ), all valid for  $Q(q; k)$ . The sum of all these inequalities is precisely the inequality (24) which proves that this inequality is redundant, and the proof is complete.  $\square$

### 5.3 Completeness

The goal of this subsection is to prove that the facet defining inequalities we have identified in the previous sections in fact give a complete linear description of  $Q(q; k)$ .

**Theorem 14** *A complete linear description of  $Q(q; k)$  is given by the set size inequalities and the  $q$ -average inequalities.*

The proof of this result is rather long, so we first sketch the main idea. Let  $R(q; k)$  denote the polyhedron defined by the set size inequalities and the  $q$ -average inequalities. Then we have  $Q(q; k) \subseteq R(q; k) \subseteq P(q; k)$ . Thus, by showing that  $R(q; k)$  is integral, the theorem will follow. In fact, we get  $Q(q; k) = Q(q; k)_I \subseteq R(q; k)_I \subseteq P(q; k)_I = Q(q; k)$ , which combined with the integrality of  $R(q; k)$  gives  $Q(q; k) = R(q; k)_I = R(q; k)$ . We shall prove that  $R(q; k)$  is integral by studying LP problems over  $R(q; k)$ . It turns out that these can be solved explicitly using the close connection to optimizing over  $P(q; k)$ . First, however, we need a couple of results for convex cones.

**Lemma 15** *Let  $C = \text{cone}(G) \subseteq \mathbb{R}^n$ , where  $G \subset \mathbb{R}^n$  is a finite set of vectors, and let  $a \in \mathbb{R}^n$ . Define  $G_0$ ,  $G_-$  and  $G_+$  as the set of elements  $g \in G$  satisfying  $a^T g = 0$ ,  $a^T g < 0$  and  $a^T g > 0$ , respectively. Assume that for each  $g_1 \in G_-$ ,  $g_2 \in G_+$  and  $\lambda_1, \lambda_2 \geq 0$  such that  $a^T(\lambda_1 g_1 + \lambda_2 g_2) = 0$ , we have that  $\lambda_1 g_1 + \lambda_2 g_2 \in \text{cone}(G_0)$ . Then  $C$  can be decomposed as  $C = \text{cone}(G_0 \cup G_-) \cup \text{cone}(G_0 \cup G_+)$ .*

**Proof.** Clearly we can decompose  $C$  into  $C = C_- \cup C_+$ , where  $C_- = \{x \in C \mid a^T x \leq 0\}$  and  $C_+ = \{x \in C \mid a^T x \geq 0\}$ . It is therefore sufficient to show that  $C_- = \text{cone}(G_0 \cup G_-)$  and  $C_+ = \text{cone}(G_0 \cup G_+)$ . From the “double description method” of Motzkin (see e.g. [9]) it follows that  $C_- = \text{cone}(G_- \cup G_0 \cup G')$  where  $G'$  is the (finite) set of vectors of the form  $\lambda_1 g_1 + \lambda_2 g_2$ , where  $g_1 \in G_-$ ,  $g_2 \in G_+$  and  $\lambda_1, \lambda_2 \geq 0$  is chosen such that  $a^T(\lambda_1 g_1 + \lambda_2 g_2) = 0$ . Now, from the assumption of the lemma, we have that  $G' \subseteq \text{cone}(G_0)$ , which implies that  $C_- = \text{cone}(G_0 \cup G_-)$  as desired. The proof of  $C_+ = \text{cone}(G_0 \cup G_+)$  is similar.  $\square$

**Lemma 16** *Let  $1 \leq s < k < n$  and let the cone  $K \subseteq \mathbb{R}^n$  be the solution set to the following homogeneous linear system:*

$$\begin{aligned} (i) \quad & z_1 \geq \dots \geq z_n \geq 0; \\ (ii) \quad & (k-s)z_s \geq \sum_{j=s+1}^n z_j; \\ (iii) \quad & (k-r)z_r \leq \sum_{j=r+1}^n z_j \quad \text{for } r = s+1, \dots, k-1. \end{aligned} \tag{25}$$

Then  $K = \text{cone}(G)$ , where  $G$  consists of the vectors (i)  $g^r$  for  $r \in \{1, \dots, s, k\}$  where  $g_j^r = 1$  when  $j \leq r$  and  $g_j^r = 0$  otherwise; and (ii)  $g^{d,h}$  for  $h = s, \dots, k-1$  and  $d = k+1, \dots, n$  where  $g_j^{d,h} = d-h$  for  $j \leq h$ ,  $g_j^{d,h} = k-h$  for  $j = h+1, \dots, d$  and  $g_j^{d,h} = 0$  otherwise.

**Proof.** Define “modified tail averages”  $\tilde{z}_{r+1:n} := (1/(k-r)) \sum_{j=r+1}^n z_j$  for  $r = s, \dots, k-1$ , and note that  $K = \{z \in \mathbb{R}^n \mid z_1 \geq \dots \geq z_n \geq 0, z_s \geq \tilde{z}_{s+1:n}, z_r \leq \tilde{z}_{r+1:n}, r = s+1, \dots, k-1\}$ . We claim that

$$K = \{z \in \mathbb{R}^n \mid z_1 \geq \dots \geq z_n \geq 0, z_s \geq \tilde{z}_{s+1:n}, z_{s+1} \leq \tilde{z}_{s+2:n}\}, \quad (26)$$

so each inequality  $z_r \leq \tilde{z}_{r+1:n}$  for  $r = s+2, \dots, k-1$  is redundant. To see this, observe first that  $\tilde{z}_{r:n} \in \text{conv}(\{z_r, \tilde{z}_{r+1:n}\})$  (similar to Lemma 1). Therefore, if  $z$  satisfies  $z_1 \geq \dots \geq z_n \geq 0, z_s \geq \tilde{z}_{s+1:n}, z_{s+1} \leq \tilde{z}_{s+2:n}$ , then it also satisfies  $z_{s+2} \leq \tilde{z}_{s+3:n}$ . (Otherwise, if  $z_{s+2} > \tilde{z}_{s+3:n}$ , we get  $z_{s+1} \geq z_{s+2} > \tilde{z}_{s+2:n} > \tilde{z}_{s+3:n}$ ; a contradiction.) We get (26) by repeating this argument for the remaining inequalities.

We shall find generators for  $K$ , or, equivalently, all the 1-dimensional faces (edges) of  $K$ . Each edge  $F$  of  $K$  is obtained by setting  $n-1$  linearly independent among the  $n+2$  inequalities  $a_i^T x \leq 0$ ,  $i = 1, \dots, n+2$  in (26) to equality. We here define  $a_i = -e_i + e_{i+1}$  for  $i = 1, \dots, n-1$ ,  $a_n = -e_n$ ,  $a_{n+1} = -(k-s)e_s + \sum_{j=s+1}^n e_j$  and  $a_{n+2} = -(k-s-1)e_{s+1} + \sum_{j=s+2}^n e_j$ .

We first determine the family  $\mathcal{S}$  consisting of all subsets  $S$  of  $\mathbb{N}_{n+2}$  with  $|S| = n-1$  such that the  $n-1$  vectors  $a_i$ ,  $i \in S$  are linearly independent. Let  $S \subset \mathbb{N}_{n+2}$  be such that  $|S| = n-1$ , and we distinguish between three different cases.

*Case 1:*  $S \subset \mathbb{N}_n$ . Then  $S \in \mathcal{S}$ , because  $a_i$ ,  $i \leq n$  are linearly independent and therefore each subset of these vectors is linearly independent.

*Case 2:*  $|S \cap \mathbb{N}_n| = n-2$ . Let  $S = \mathbb{N}_n \setminus \{i_1, i_2\}$ . If  $n+1 \in S$ , then  $S \in \mathcal{S}$  iff  $a_{n+1}$  is not a linear combination of  $a_i$ ,  $i \leq n$ ,  $i \neq i_1, i_2$ . But the unique solution  $y$  of  $\sum_{i=1}^n y_i a_i = a_{n+1}$  is  $y_i = 0$  for  $i \leq s-1$ ,  $y_i = k-i$  for  $i = s+1, \dots, n$ . Consequently,  $S \in \mathcal{S}$  iff either  $i_1$  or  $i_2$  lies outside  $\{1, \dots, s-1, k\}$ . If, alternatively,  $n+2 \in S$ , similar arguments show that  $S \in \mathcal{S}$  iff either  $i_1$  or  $i_2$  lies outside  $\{1, \dots, s, k\}$ .

*Case 3:*  $n+1, n+2 \in S$ . Let  $\{i_1, i_2, i_3\} = \mathbb{N}_n \setminus S$ . By direct elimination one can show that  $\sum_{i=1}^{n+2} y_i a_i = 0$  iff  $y$  satisfies  $y_1 = \dots = y_{s-1} = 0$ ,  $y_s = -(k-s)y_{n+1}$  and  $y_j = -(k-j)(y_{n+1} + y_{n+2})$  for  $j = s+1, \dots, n$ . It follows that  $S \in \mathcal{S}$  iff  $s \in \{i_1, i_2, i_3\}$  and  $\{i_1, i_2, i_3\} \cap (\{s+1, \dots, n\} \setminus \{k\})$  is nonempty.

It is now possible to determine the edges of  $K$  as follows. For each  $S \in \mathcal{S}$  one must find the solution set to the system  $a_i^T x = 0$  for  $i \in S$ ,  $a_i^T x \leq 0$  for  $i \in \mathbb{N}_{n+2} \setminus S$ . This can be done by a rather long, but elementary discussion which is omitted here. We just remark that the generators  $g^r$  for  $r \in \{1, \dots, s, k\}$  are found in Case 1 above, while the generators  $g^{d,h}$  for  $h = s, \dots, k-1$  and  $d = k+1, \dots, n$  are found from Case 2. Finally, Case 3 does not lead to any other generators than those already found. This proves that  $K = \text{cone}(G)$  as desired.  $\square$

**Proposition 17** *Let  $s \in \{0, \dots, k-1\}$  and define  $\bar{m} = \lfloor \bar{q}_{s:k} \rfloor$  and  $\hat{m} = \bar{m} + 1$ . Assume that the tail average  $\bar{q}_{s:k}$  is fractional. Then the  $q$ -averages  $q^{\bar{m}}$  and  $q^{\hat{m}}$  are adjacent vertices of  $R(q; k)$  and the edge  $F$  of  $R(q; k)$  between these vertices is the solution set of the linear system*

$$\begin{aligned}
(i) \quad & x(\mathbb{N}_r) = q(\mathbb{N}_r) && \text{for } r = 1, \dots, s-1; \\
(ii) \quad & x(\mathbb{N}_s \cup S) = Q && \text{for } S \subseteq \{s+1, \dots, n\} \text{ with} \\
& && |S| = k-s, s+1 \in S; \\
(iii) \quad & a^{s,t}x = \alpha^{s,t} && \text{for } k < t \leq n; \\
(iv) \quad & x(\mathbb{N}_s) \leq q(\mathbb{N}_s); \\
(v) \quad & x(\mathbb{N}_{k+1} \setminus \{s+1\}) \leq Q.
\end{aligned} \tag{27}$$

The vertex  $q^{\bar{m}}$  (resp.  $q^{\hat{m}}$ ) is the unique solution of (27)(i)-(iii) and (iv) (resp. (v)) set to equality.

**Proof.** We first remark that all the equations in (27)(i)-(iii) are obtained by setting valid inequalities for  $R(q; k)$  (set size and  $q$ -average inequalities) to equality. Thus the solution set  $F$  of the linear system (27) is a face of  $R(q; k)$ , and we shall determine this set.

Assume that  $x$  satisfies (27). From (27)(i) it follows that  $x_j = q_j$  for  $j = 1, \dots, s-1$ . Combining this with (27)(ii) we get  $x(S) = Q - x_s - x_{s+1} - \sum_{j=1}^{s-1} q_j$  for all  $S \subseteq \{s+2, \dots, n\}$ ,  $|S| = k-s-1$ . In particular, for  $i, j \in \{s+2, \dots, n\}$  and some  $S' \subseteq \{s+2, \dots, n\} \setminus \{i, j\}$  with  $|S'| = k-s-2$  we get  $x(S') + x_i = Q - x_s - x_{s+1} - \sum_{j=1}^{s-1} q_j = x(S') + x_j$  and therefore  $x_i = x_j$  for all  $i, j \geq s+2$ , i.e.  $x_i = x_{s+2}$  for all  $i \geq s+2$ . Thus from (27)(ii) and (iii) we obtain  $x_s + x_{s+1} + (k-s-1)x_{s+2} = Q - \sum_{j=1}^{s-1} q_j = \sum_{j=s}^k q_j$  and  $(n-s-\delta^s)(\sum_{j=1}^{s-1} q_j + x_s) + (k-s-\delta^s)(x_{s+1} + (n-s-1)x_{s+2}) = \alpha^{s,n}$  where  $\alpha^{s,n} = (n-s-\delta^s)(\sum_{j=1}^{s-1} q_j + \Delta) + (k-s-\delta^s)(n-s)\bar{m}$ ,  $\Delta = \sum_{j=s}^k q_j - (k-s)\bar{m}$ . Note that the coefficient vectors of these two equations are linearly independent and from the first of these we get  $x_{s+1} = \sum_{j=s}^k q_j - x_s - (k-s-1)x_{s+2}$ . Substituting this expression into the second equation, gives (after some calculations) that  $x = x(\beta)$  for some  $\beta \in \mathbb{R}$  where we define  $x(\beta)_j = q_j$  for  $j \leq s-1$ ,  $x(\beta)_s = \sum_{j=s}^k q_j - \delta^s(\bar{m}+1) - (k-s-\delta^s)\beta$ ,  $x(\beta)_{s+1} = \delta^s(\bar{m}+1) + (1-\delta^s)\beta$ ,  $x(\beta)_{s+2} = \dots = x_n = \beta$ . This is a parametric form of the solution set of (27)(i)-(iii).

We claim that

$$F = \{x(\beta) \mid \bar{m} \leq \beta \leq \hat{m}\}. \tag{28}$$

Assume first that  $x \in F$ , and by the discussion above  $x = x(\beta)$  for some real  $\beta$ , i.e.  $x(\beta) = (q_1, \dots, q_{s-1}, x(\beta)_s, x(\beta)_{s+1}, \beta, \dots, \beta)$  where  $x(\beta)_s = \sum_{j=s}^k q_j - \delta^s \hat{m} - (k-s-\delta^s)\beta$ , and  $x(\beta)_{s+1} = \delta^s \hat{m} + (1-\delta^s)\beta$ . If  $\beta < \bar{m}$ , we get  $x(\beta)_s = \sum_{j=s}^k q_j - \delta^s \hat{m} - (k-s-\delta^s)\beta > \sum_{j=s}^k q_j - \delta^s \hat{m} - (k-s-\delta^s)\bar{m} = \sum_{j=s}^k q_j - (k-s)\bar{m} - \delta^s = q_s$  i.e.  $x(\beta)_s > q_s$  which gives  $\sum_{j=1}^s x(\beta)_j > q(\mathbb{N}_s)$ ; a contradiction. (Here the last strict inequality was due to  $\hat{m} > \bar{q}_{s+1:k} = (1/(k-s)) \sum_{j=s+1}^k q_j$ ). This proves that  $\beta \geq \bar{m}$ .

Next,  $\sum_{j \leq k+1, j \neq s+1}^k x(\beta)_j = \sum_{j=1}^{s-1} q_j + (\sum_{j=s}^k q_j - \delta^s \hat{m} - (k-s-\delta^s)\beta) + (k-s)\beta = Q + \delta^s(\beta - \hat{m})$ . But  $\sum_{j \leq k+1, j \neq s+1}^k x_j \leq Q$  is a set size inequality (and therefore valid for  $R(q; k)$ ), so we see that  $x(\beta) \in R(q; k)$  implies that  $\beta \leq \hat{m}$  since  $\delta^s > 0$  because we assume that  $\bar{q}_{s+1:k}$  is fractional. This proves that  $F \subseteq \{x(\beta) \mid \bar{m} \leq \beta \leq \hat{m}\}$ . The opposite inclusion is easy to prove by direct calculations, and therefore the representation (28) for  $F$  holds. This proves that  $F$  is an edge of  $R(q; k)$ , and that its two vertices are  $x(\bar{m}) = q^{\bar{m}}$  and  $x(\hat{m}) = q^{\hat{m}}$ . Furthermore, the description of these vertices stated in the last part of the proposition also holds; this follows from the arguments in our proof of (28).  $\square$

**Proof of Theorem 14:** From the comments in the beginning of this subsection, it suffices to prove that  $R(q; k)$  is integral. We do this by proving that the LP problem (R)  $\max \{c^T x \mid x \in R(q; k)\}$  has an optimal solution which is integral. Let  $c \in \mathbb{R}^n$  and we may assume that  $c_1 \geq \dots \geq c_n \geq 0$  (otherwise a suitable permutation of the solution vector would be required). From Theorem 2 it follows that an optimal solution of the LP problem (P)  $\max \{c^T x \mid x \in P(q; k)\}$  is the  $q$ -average  $w^s$  where  $s \in \{0, \dots, k-1\}$  is chosen such that  $c_s \geq (1/(k-s)) \sum_{j=s+1}^n c_j$  and  $c_r \leq (1/(k-r)) \sum_{j=r+1}^n c_j$  for  $r = s+1, \dots, k-1$ . Thus  $c \in K_s$  where we let  $K_s$  be the cone  $K$  of Lemma 16. Let  $\bar{m} = \lfloor \bar{q}_{s+1:k} \rfloor$  and  $\hat{m} = \bar{m} + 1$ , and note that  $q^{\bar{m}}, q^{\hat{m}} \in Q(q; k) \subseteq R(q; k)$ . The goal is to prove using linear programming duality that *either  $q^{\bar{m}}$  or  $q^{\hat{m}}$  is an optimal solution of the problem (R)*. First, however, we note that if  $\bar{q}_{s:k}$  is integral, then  $q^{\bar{m}} = w^s$  is an optimal solution in (R) (as it is integral and optimal in a relaxation). Thus we consider the case with  $\bar{q}_{s:k}$  fractional in the discussion below.

Define  $a := q^{\hat{m}} - q^{\bar{m}}$ , so  $a_j = 0$  for  $j \leq s-1$ ,  $a_s = -(k-s-\delta^s)$ ,  $a_{s+1} = 1 - \delta^s$ , and  $a_j = 1$  for  $j \geq s+2$  (see (20) for the definition of  $\delta^s$ ). Note that  $a$  is a direction vector for the edge  $F$  of  $R(q; k)$  between the vertices  $q^{\bar{m}}$  and  $q^{\hat{m}}$ . Let  $Z$  denote the set of coefficient vectors from the equations (27)(1)-(iii), namely  $\chi^{\mathbb{N}_r}$  for  $r = 1, \dots, s-1$ ;  $\chi^{\mathbb{N}_s \cup S}$  for  $S \subseteq \{s+1, \dots, n\}$ ,  $|S| = k-s$ ,  $s+1 \in S$ ; and  $a^{s,t}$  for  $k < t \leq n$ . It follows from Proposition 17 that  $a$  and  $z$  are orthogonal for each  $z \in Z$ , i.e.,

$$a^T z = 0 \text{ for all } z \in Z. \quad (29)$$

We shall study the cone  $K_s$  further. Let  $G$  be the generators for  $K_s$  known from Lemma 16, and define  $G_0$ ,  $G_-$  and  $G_+$  as the set of elements  $g \in G$  satisfying  $a^T g = 0$ ,  $a^T g < 0$  and  $a^T g > 0$ , respectively.

*Claim 1:*  $G_0 = \{g^r \mid r = 1, \dots, s-1, k\} \cup \{g^{d,h} \mid h = s+1, \dots, k-1; d = k+1, \dots, n\}$ ,  $G_- = \{g^s\}$  and  $G_+ = \{g^{d,s} \mid d = k+1, \dots, n\}$ . For  $r \leq s-1$ , we have  $g^r \in Z$  so (29) gives  $a^T g^r = 0$ , and  $g^r \in G_0$ . Furthermore,  $a^T g^s = -(k-s-\delta^s) < 0$  so  $g^s \in G_-$ , and  $g^k = \chi^{\mathbb{N}_s} + \chi^S \in Z$  for  $S = \{s+1, \dots, k\}$  so  $g^k \in G_0$ . Consider  $g^{d,h}$  for  $h \in \{s+1, \dots, k-1\}$  and  $d \in \{k+1, \dots, n\}$ . Define subsets  $S_1, \dots, S_{d-h}$  of  $\{h+1, \dots, d\}$  such that the  $(d-h)$ -dimensional square matrix with  $i$ 'th row equal to the incidence vector of  $S_i$  viewed as a subset of  $\{h+1, \dots, d\}$  equals the



0/1-circulant matrix  $C_{k-h, d-h}$ . Since  $h \geq s+1$  we have  $\chi^{\mathbb{N}_h} + \chi^{S_i} \in Z$  for each  $i$ , and the sum of these vectors is  $g^{d,h}$  which proves that  $g^{d,h} \in G_0$ . Finally, we consider  $g^{d,h}$  for  $h = s$  and  $d \in \{k+1, \dots, n\}$ . We then have  $a^T g^{d,s} = -(d-s)(k-s-\delta^s) + (k-s)(1-\delta^s) + \sum_{j=s+2}^d (k-s) = \delta^s(d-k) > 0$  and therefore  $g^{d,s} \in G_+$  for  $d \in \{k+1, \dots, n\}$ . This completes the proof of Claim 1.

*Claim 2:*  $K_s = \text{cone}(G_0 \cup G_-) \cup \text{cone}(G_0 \cup G_+)$ . To prove this we shall apply Lemma 15 with  $K = K_s = \text{cone}(G)$  and  $a$  as above, and we need to verify the hypothesis of that lemma. So let  $g_1 \in G_-$ ,  $g_2 \in G_+$  and  $\lambda_1, \lambda_2 \geq 0$  be such that  $a^T \gamma = 0$  where  $\gamma = \lambda_1 g_1 + \lambda_2 g_2$ . It follows from Claim 1 that  $g_1 = g^s$  and  $g_2 = g^{d,s}$  for some  $d \in \{k+1, \dots, n\}$ . Therefore  $\gamma_j = \lambda_1 + \lambda_2(d-s)$  for  $j \leq s$ ,  $\gamma_j = \lambda_2(k-s)$  for  $s+1 \leq j \leq d$  and  $\gamma_j = 0$  for  $j > d$ , which leads to  $a^T \gamma = -(k-s-\delta^s)(\lambda_1 + \lambda_2(d-s)) + (1-\delta^s)\lambda_2(k-s) + \sum_{j=s+2}^d \lambda_2(k-s) = -(k-s-\delta^s)\lambda_1 - (k-s-\delta^s)\lambda_2(d-s) + \lambda_2(k-s)(d-s-\delta^s) = -(k-s-\delta^s)\lambda_1 + \lambda_2\delta^s(d-k)$ . Since  $a^T \gamma = 0$  we get  $\lambda_2\delta^s(d-k) = (k-s-\delta^s)\lambda_1$ . This can be used in elementary calculations to prove that  $\gamma = \mu a^{s,d}$  for some  $\mu > 0$  where  $(a^{s,d})^T x \leq \alpha^{s,d}$  is a  $q$ -average inequality (see (20)). Since  $a^{s,d} \in Z$ , we get from (29) that  $a^T \gamma = 0$ , which verifies that Lemma 15 can be applied, and Claim 2 follows.

We next relate the generators in  $G$  to defining inequalities for  $R(q; k)$  that are active in either  $q^{\bar{m}}$  or  $q^{\bar{n}}$ . Let us say that  $g \in \mathbb{R}^n$  is  $R$ -active in a vertex  $z$  of  $R(q; k)$  if there is a linear system  $Ax \leq b$  consisting of defining inequalities for  $R(q; k)$  (i.e., set size and  $q$ -average inequalities) such that  $Az = b$  and  $y^T A = g^T$  for some  $y \geq 0$ .

*Claim 3:* Each  $g \in G_0$  is  $R$ -active in both  $q^{\bar{m}}$  and  $q^{\bar{n}}$ . Furthermore,  $g^s$  is  $R$ -active in  $q^{\bar{m}}$  and each  $g \in G_+$  is  $R$ -active in  $q^{\bar{n}}$ . We first consider  $g \in G_0$  (see Claim 1) and use the results of Proposition 17. For  $r \leq s-1$  or  $r = k$  the set size inequality  $x(\mathbb{N}_r) \leq q(\mathbb{N}_r)$  is active in both  $q^{\bar{m}}$  and  $q^{\bar{n}}$ , and since  $(g^r)^T x = x(\mathbb{N}_r)$ , we see that  $g^r$  is  $R$ -active in these two vertices. Consider  $g^{d,h}$  for  $h \in \{s+1, \dots, k-1\}$ ,  $d \in \{k+1, \dots, n\}$ . From the proof of Claim 1, we have  $\sum_i (\chi^{\mathbb{N}_h} + \chi^{S_i}) = g^{d,h}$ . But each set size inequality  $x(\mathbb{N}_h \cup S_i) \leq Q$  is active in both  $q^{\bar{m}}$  and  $q^{\bar{n}}$ , so  $(g^{d,h})^T x \leq (d-h)Q$  is a sum of active set size inequalities in these two vertices which implies that  $g^{d,h}$  is  $R$ -active. The inequality  $x(\mathbb{N}_s) \leq q(\mathbb{N}_s)$  is active in  $q^{\bar{m}}$  which implies that  $g^s$  is  $R$ -active in  $q^{\bar{m}}$ . Finally, consider  $g^{d,s}$  for  $d \in \{k+1, \dots, n\}$ . As in the case with  $h > s$  we can write  $g^{d,s} = \sum_i (\chi^{\mathbb{N}_s} + \chi^{S_i})$  with subsets  $S_i$  defined analogously. Now, each of the inequalities  $x(\mathbb{N}_s \cup S_i) \leq Q$  is active in  $q^{\bar{n}}$ , and therefore  $g^{d,s}$  is  $R$ -active in  $q^{\bar{n}}$  and Claim 3 has been proved.

We can now put all these results together. Let  $Ax \leq b$  be all the set size and  $q$ -average inequalities, so  $R(q; k) = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ . Recall that the objective function  $c$  satisfies  $c \in K_s$ , so by Claim 2 we have either (i)  $c \in \text{cone}(G_0 \cup G_-)$  or (ii)  $c \in \text{cone}(G_0 \cup G_+)$ . Consider first case (i); we shall use LP duality to prove that  $q^{\bar{m}}$  is an optimal solution of the problem (R)  $\max \{c^T x \mid Ax \leq b\}$ . Since  $c \in \text{cone}(G_0 \cup G_-)$ , we can write  $c$  as a conical combination of generators in  $G_0 \cup G_-$ . But all these

generators are  $R$ -active in  $q^{\bar{m}}$  (see Claim 3), which proves that there is an  $y \geq 0$  such that  $y^T A = c^T$  and with  $a_i^T q^{\bar{m}} = b_i$  whenever  $y_i > 0$ . But then  $q^{\bar{m}}$  is primal feasible (in (R)),  $y$  is dual feasible and complementary slackness holds, so by linear programming duality  $q^{\bar{m}}$  is optimal in (R) and  $y$  is optimal in the dual problem. In case (ii), when  $c \in \text{cone}(G_0 \cup G_+)$ , similar arguments prove that  $q^{\bar{m}}$  is optimal in (R). Thus we have shown that the problem (R) always has an optimal integral solution, meaning that  $R(q; k)$  is an integral polyhedron and the proof is complete.  $\square$

**Example.** For our example with  $k = 3$ ,  $n = 5$  and  $q = (7, 2, 1)$  a complete and nonredundant linear description of  $Q(q; k)$  consists of all “permutations” of these inequalities: the set size inequalities  $x_1 \leq 7$ ,  $x(\mathbb{N}_2) \leq 9$ ,  $x(\mathbb{N}_3) \leq 10$ ; the extended set size inequalities  $x(\mathbb{N}_4) \leq 13$  and  $x(\mathbb{N}_5) \leq 16$ , and finally the other  $q$ -average inequalities  $2x_1 + x_2 + x_3 + x_4 \leq 18$  and  $3x_1 + x_2 + x_3 + x_4 + x_5 \leq 26$ .

An algorithm for problem the problem  $\max \{c^T x \mid x \in Q(q; k)\}$  can also be given. First, solve the problem  $\max \{c^T x \mid x \in P(q; k)\}$  and let  $w^s$  be an optimal solution (see section 3). Then either  $q^{\bar{m}}$  or  $q^{\hat{m}}$  (where  $\bar{m} = \lfloor q_{s+1:k} \rfloor$  and  $\hat{m} = \bar{m} + 1$ ) must be optimal, and an optimal solution is therefore found by comparing the associated two objective function values.

## 6 Conclusions

We have introduced the concept of weak  $k$ -majorization and studied associated polyhedra. Complete inner and outer descriptions were found for the  $k$ -majorization polyhedron  $P(q; k)$  consisting of all vectors weakly  $k$ -majorized by a given vector as well as for the integer hull  $Q(q; k)$  of  $P(q; k)$ . Efficient algorithms for solving linear programming and integer linear programming problems over  $P(q; k)$  were also given.

An interesting direction for further work is to study well-known problems in network design or network flows with additional  $k$ -majorization constraints. Both structural and algorithmic results would be of interest, and some work in this direction is ongoing. Also other problems in operations research for which diversification is of interest may be attacked similarly.

## References

- [1] Björner, A. et al., *Theory of Oriented Matroids*, Cambridge Univ. Press, Cambridge, 1993.
- [2] Dahl, G., *Pseudo experiments and majorization*. Thesis, University of Oslo, 1984.
- [3] Edmonds, J., Matroids and the greedy algorithm, *Mathematical Programming* 1 (1971) 127-136.

- [4] Folkman, J.H. and Fulkerson, D.R., "Edge colorings in bipartite graphs," In *Combinatorial mathematics and its applications* (R.C. Bose and T.A. Dowling, eds.), Chapter 31, pp. 561-577, Univ. of North Carolina, Chapel Hill, 1969.
- [5] Fujishige, S. *Submodular functions and optimization*, Annals of discrete mathematics, Vol. 47. Amsterdam, North Holland, 1991.
- [6] Grötschel, M., Lovász, L. and Schrijver, A., *Geometric algorithms and combinatorial optimization*, Springer, 1988.
- [7] Hardy, G.H, Littlewood, J.E. and Polya, G., *Inequalities* (2. ed.), Cambridge Mathematical Library, Cambridge University Press, 1988.
- [8] Marshall, A.W. and Olkin, I., *Inequalities: Theory of Majorization and Its Applications*, New York, Academic Press, 1979.
- [9] Pulleyblank, W.R., "Polyhedral combinatorics," in Handbooks in Operations Research, Vol. 1, *Optimization*, ed. Nemhauser et al., North-Holland, 1989.
- [10] Schrijver, A., *Theory of linear and integer programming*, Wiley, 1986.
- [11] Torgersen, E., *Comparison of Statistical Experiments*, Cambridge University Press, Cambridge, 1992.
- [12] Ziegler, G., *Lectures on polytopes*, Konrad-Zuse-Zentrum für Informationstechnik Berlin, Technical Report TF 93-6, 1993.